

**PCT** WELTORGANISATION FÜR GEISTIGES EIGENTUM  
Internationales Büro  
INTERNATIONALE ANMELDUNG VERÖFFENTLICH NACH DEM VERTRAG ÜBER DIE  
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)



<b>(51) Internationale Patentklassifikation 6 :</b> <b>C12N 15/00</b>		<b>A2</b>	<b>(11) Internationale Veröffentlichungsnummer:</b> <b>WO 99/54446</b> <b>(43) Internationales Veröffentlichungsdatum:</b> <b>28. Oktober 1999 (28.10.99)</b>
<b>(21) Internationales Aktenzeichen:</b> <b>PCT/DE99/01096</b> <b>(22) Internationales Anmeldedatum:</b> <b>8. April 1999 (08.04.99)</b>		<b>(81) Bestimmungsstaaten:</b> JP, US, europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
<b>(30) Prioritätsdaten:</b> 198 18 598.7 19. April 1998 (19.04.98) DE		<b>Veröffentlicht</b> <i>Ohne internationalem Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</i>	
<b>(71) Anmelder (für alle Bestimmungsstaaten ausser US):</b> META-GEN GESELLSCHAFT FÜR GENOMFORSCHUNG MBH [DE/DE]; Ihnestrasse 63, D-14195 Berlin (DE).			
<b>(72) Erfinder; und</b> <b>(75) Erfinder/Anmelder (nur für US):</b> SPECHT, Thomas [DE/DE]; Grabenstrasse 14, D-12209 Berlin (DE). HINZMANN, Bernd [DE/DE]; Parkstrasse 19, D-13127 Berlin (DE). SCHMITT, Armin [DE/DE]; Laubacher Strasse 6/II, D-14197 Berlin (DE). PILARSKY, Christian [DE/DE]; Heinrich-Lange-Strasse 13c, D-01474 Schönefeld-Weißenberg (DE). DAHL, Edgar [DE/DE]; Eleonore-Procheska-Strasse 6, D-14480 Potsdam (DE). ROSENTHAL, André [DE/DE]; Koppenplatz 10, D-10115 Berlin (DE).			

**(54) Title:** HUMAN NUCLEIC ACID SEQUENCES FROM NORMAL PANCREAS TISSUE

**(54) Bezeichnung:** MENSCHLICHE NUKLEINSÄURESEQUENZEN AUS PANKREASNORMALGEWEBE

**(57) Abstract**

The invention relates to human nucleic acid sequences – mRNA, cDNA, genomic sequences – from normal pancreas tissue, coding for the genetic product or parts thereof. The invention also relates to the utilization of said sequences, to the polypeptides obtained through said sequences and their utilization.

**(57) Zusammenfassung**

Es werden menschliche Nukleinsäuresequenzen – mRNA, cDNA, genomicsche Sequenzen – aus Pankreasnormalgewebe, die für Genprodukte oder Teile davon kodieren, und deren Verwendung beschrieben. Es werden weiterhin die über die Sequenzen erhältlichen Polypeptide und deren Verwendung beschrieben.

**LEDIGLICH ZUR INFORMATION**

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Leitland	SZ	Swasiland
AZ	Aserbaidschan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland	ML	Mali	TR	Türkei
BG	Bulgarien	HU	Ungarn	MN	Mongolei	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MR	Mauretanien	UA	Ukraine
BR	Brasilien	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Island	MX	Mexiko	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	NE	Niger	UZ	Usbekistan
CF	Zentralafrikanische Republik	JP	Japan	NL	Niederlande	VN	Vietnam
CG	Kongo	KE	Kenia	NO	Norwegen	YU	Jugoslawien
CH	Schweiz	KG	Kirgisistan	NZ	Neuseeland	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	PL	Polen		
CM	Kamerun	KR	Republik Korea	PT	Portugal		
CN	China	KZ	Kasachstan	RO	Rumänien		
CU	Kuba	LC	St. Lucia	RU	Russische Föderation		
CZ	Tschechische Republik	LI	Liechtenstein	SD	Sudan		
DE	Deutschland	LK	Sri Lanka	SE	Schweden		
DK	Dänemark	LR	Liberia	SG	Singapur		
EE	Estland						

**Menschlich Nukleinsäuresequenzen aus Pankreasnormalg webe**

5 Die Erfindung betrifft menschliche Nukleinsäuresequenzen aus Pankreasnormalgewebe, die für Genprodukte oder Teile davon kodieren, deren funktionale Gene, die mindestens ein biologisch aktives Polypeptid kodieren und deren Verwendung.

10 Die Erfindung betrifft weiterhin die über die Sequenzen erhältlichen Polypeptide und deren Verwendung.

15 Eine der Krebstodesursachen ist der Pankreastumor, für dessen Bekämpfung neue Therapien notwendig sind. Bisher verwendete Therapien, wie z.B. Chemotherapie, Hormontherapie oder chirurgische Entfernung des Tumorgewebes, führen häufig nicht zu einer vollständigen Heilung.

20 15 Das Phänomen Krebs geht häufig einher mit der Über- oder Unterexpression gewisser Gene in den entarteten Zellen, wobei noch unklar ist, ob diese veränderten Expressionsraten Ursache oder Folge der malignen Transformation sind. Die Identifikation solcher Gene wäre ein wesentlicher Schritt für die Entwicklung neuer Therapien gegen Krebs. Der spontanen Entstehung von Krebs geht häufig eine Vielzahl von Mutationen voraus. Diese können verschiedenste Auswirkungen auf das Expressionsmuster in dem betroffenen Gewebe haben, wie z.B. Unter- oder Überexpression, aber auch Expression verkürzter Gene. Mehrere solcher Veränderungen durch solche Mutationskaskaden können schließlich zu bösartigen 25 Entartungen führen. Die Komplexität solcher Zusammenhänge erschwert die experimentelle Herangehensweise sehr.

30 Für die Suche nach Kandidatengenen, d.h. Genen, die im Vergleich zum Tumorgewebe im normalen Gewebe stärker exprimiert werden, wird eine Datenbank verwendet, die aus sogenannten ESTs besteht. ESTs (Expressed Sequence Tags) sind Sequenzen von cDNAs, d.h. revers transkribierten mRNAs, den Molekülen also, die die Expression von Genen widerspiegeln. Die EST-Sequenzen werden für normale und entartete Gewebe ermittelt. Solche Datenbanken werden von verschiedenen Betreibern z.T. kommerziell angeboten. Die ESTs der LifeSeq-35 Datenbank, die hier verwendet wird, sind in der Regel zwischen 150 und 350 Nukleotide lang. Sie repräsentieren ein für ein bestimmtes Gen unverkennbares Muster, obwohl dieses Gen normalerweise sehr viel länger ist (> 2000 Nukleotide). Durch Vergleich der Expressionsmuster von normalen und Tumorgewebe können ESTs identifiziert werden, die für die Tumorentstehung und -proliferation wichtig sind.

40 Es besteht jedoch folgendes Problem: Da durch unterschiedliche Konstruktionen der cDNA-Bibliotheken die gefundenen EST-Sequenzen zu unterschiedlichen Regionen eines unbekannten Gens gehören können, ergäbe sich in einem solchen Fall ein völlig falsches Verhältnis des Vorkommens dieser ESTs in dem jeweiligen Gewebe. Dieses würde erst bemerkt werden, wenn das vollständige Gen bekannt ist und somit 45 die ESTs dem gleichen Gen zugeordnet werden können.

50 Es wurde nun gefunden, daß diese Fehlermöglichkeit verringert werden kann, wenn zuvor sämtliche ESTs aus dem jeweiligen Gewebstyp assembliert werden, bevor die Expressionsmuster miteinander verglichen werden. Es wurden also überlappende ESTs ein und desselben Gens zu längeren Sequenzen zusammengefaßt (s. Fig. 1, Fig. 2a und Fig. 3). Durch diese Verlängerung und damit Abdeckung eines wesentlich größeren Genbereichs in jeder der jeweiligen Banken sollte der oben beschriebene Fehler weitgehend vermieden werden. Da es hierzu keine bestehenden

Softwarereprodukte gab, wurden Programme für das Assemblieren von genomischen Abschnitten verwendet, die abgewandelt eingesetzt und durch eigene Programme ergänzt wurden. Ein Flowchart der Assemblierungsprozedur ist in Fig. 2b1 – 2b4 dargestellt.

5 Es konnten nun die Nukleinsäure-Sequenzen Seq. ID No 2-37, 67 gefunden werden, die als Kandidatengene beim Pankreastumor eine Rolle spielen.

10 Von besonderem Interesse sind die Nukleinsäure-Sequenzen Seq. ID No 14, 24, 25, 27-31, 35-37, 67.

15 Die Erfindung betrifft somit Nukleinsäure-Sequenzen, die ein Genprodukt oder ein Teil davon kodieren, umfassend

20 15 a) eine Nukleinsäure-Sequenz, ausgewählt aus der Gruppe der Nukleinsäure-Sequenzen Seq. ID No 14, 24, 25, 27-31, 35-37, 67.

25 20 b) eine allelische Variation der unter a) genannten Nukleinsäure-Sequenzen

oder

30 25 c) eine Nukleinsäure-Sequenz, die komplementär zu den unter a) oder b) genannten Nukleinsäure-Sequenzen ist.

35 30 Die Erfindung betrifft weiterhin eine Nukleinsäure-Sequenz gemäß einer der Sequenzen Seq. ID No 14, 24, 25, 27-31, 35-37, 67 oder eine komplementäre oder allelische Variante davon und die Nukleinsäure-Sequenzen davon, die eine 90%ige bis 95%ige Homologie zu einer humanen Nukleinsäure-Sequenz aufweisen.

40 35 Die Erfindung betrifft auch die Nukleinsäure-Sequenzen Seq. ID No 2-37, 67, die im Pankreasnormalgewebe erhöht exprimiert sind.

45 Die Erfindung betrifft ferner Nukleinsäure-Sequenzen, umfassend einen Teil der oben genannten Nukleinsäure-Sequenzen, in solch einer ausreichenden Größe, daß sie mit den Sequenzen Seq. ID No 14, 24, 25, 27-31, 35-37, 67 hybridisieren.

50 40 Die erfindungsgemäßen Nukleinsäure-Sequenzen weisen im allgemeinen eine Länge von mindestens 50 bis 4500 bp, vorzugsweise eine Länge von mindestens 150 bis 4000 bp, insbesondere eine Länge von 450 bis 3500 bp auf.

45 55 Mit den erfindungsgemäßen Teilsequenzen Seq. ID No 2-37, 67 können gemäß gängiger Verfahrenspraxis auch Expressionskassetten konstruiert werden, wobei auf der Kassette mindestens eine der erfindungsgemäßen Nukleinsäure-Sequenzen zusammen mit mindestens einer dem Fachmann allgemein bekannten Kontroll- oder regulatorischen Sequenz, wie z. B. einem geeigneten Promotor, kombiniert wird. Die erfindungsgemäßen Sequenzen können in sense oder antisense Orientierung eingefügt sein.

In der Literatur sind eine große Anzahl von Expressionskassetten bzw. Vektoren und Promotoren bekannt, die verwendet werden können.

5 Unter Expressionskassetten bzw. Vektoren sind zu verstehen: 1. bakterielle, wie z. B., phagescript, pBs,  $\phi$ X174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene), pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia), 2. eukaryontische, wie z. B. pWLneo, pSV2cat, pOG44, pXT1, pSG (Stratagene), pSVK3, pBPV, pMSG, pSVL (Pharmacia).

10 Unter Kontroll- oder regulatorischer Sequenz sind geignete Promotoren zu verstehen. Hierbei sind zwei bevorzugte Vektoren der pKK232-8 und der PCM7 Vektor. Im einzelnen sind folgende Promotoren gemeint: lacI, lacZ, T3, T7, gpt, lambda PR, trc, CMV, HSV Thymidin-Kinase, SV40, LTRs aus Retrovirus und Maus 15 Metallothionein-I.

Die auf der Expressionskassette befindlichen DNA-Sequenzen können ein Fusionsprotein kodieren, das ein bekanntes Protein und ein biologisch aktives Polypeptid-Fragment umfaßt.

20 Die Expressionskassetten sind ebenfalls Gegenstand der vorliegenden Erfindung.

Die erfindungsgemäßen Nukleinsäure-Fragmente können zur Herstellung von Vollängen-Genen verwendet werden. Die erhältlichen Gene sind ebenfalls 25 Gegenstand der vorliegenden Erfindung.

Die Erfindung betrifft auch die Verwendung der erfindungsgemäßen Nukleinsäure-Sequenzen, sowie die aus der Verwendung erhältlichen Gen-Fragmente.

30 Die erfindungsgemäßen Nukleinsäure-Sequenzen können mit geeigneten Vektoren in Wirtszellen gebracht werden, in denen als heterologer Teil die auf den Nukleinsäure-Fragmenten enthaltene genetischen Information befindet, die exprimiert wird.

35 Die die Nukleinsäure-Fragmente enthaltenden Wirtszellen sind ebenfalls Gegenstand der vorliegenden Erfindung.

Geeignete Wirtszellen sind z. B. prokaryontische Zellsysteme wie E. coli oder eukaryontische Zellsysteme wie tierische oder humane Zellen oder Hefen.

40 Die erfindungsgemäßen Nukleinsäure-Sequenzen können in sense oder antisense Form verwendet werden.

45 Die Herstellung der Polypeptide oder deren Fragment erfolgt durch Kultivierung der Wirtszellen gemäß gängiger Kultivierungsmethoden und anschließender Isolierung und Aufreinigung der Peptide bzw. Fragmente, ebenfalls mittels gängiger Verfahren. Die Erfindung betrifft ferner Nukleinsäure-Sequenzen, die mindestens eine Teilsequenz eines biologisch aktiven Polypeptids kodieren.

Ferner betrifft die vorliegende Erfindung Polypeptid-Teilsequenzen, sogenannte ORF (open-reading-frame)-Peptide, gemäß den Sequenzprotokollen Seq. ID No 39-63, 68-71.

5 Die Erfindung betrifft ferner die Polypeptid-Sequenzen, die mindestens eine 80%ige Homologie, insbesondere eine 90%ige Homologie zu den erfindungsgemäßen Polypeptid-Teilsequenzen der Seq. ID No 39-63, 68-71 aufweisen.

10 Die Erfindung betrifft auch Antikörper, die gegen ein Polypeptid oder Fragment davon gerichtet sind, welche von den erfindungsgemäßen Nukleinsäuren der Sequenzen Seq. ID No 2-37, 67 kodiert werden.

Unter Antikörper sind insbesondere monoklonale Antikörper zu verstehen.

15 Die erfindungsgemäßen Antikörper können u.a. durch ein Phage Display Verfahren identifiziert werden. Auch diese Antikörper sind Gegenstand der Erfindung.

20 Die erfindungsgemäßen Polypeptid-Teilsequenzen können in einem Phage Display Verfahren verwendet werden. Die mit diesem Verfahren identifizierten Polypeptide, die an die erfindungsgemäßen Polypeptid-Teilsequenzen binden, sind auch Gegenstand der Erfindung.

25 Ebenso können die erfindungsgemäßen Nukleinsäure-Sequenzen in einem Phage Display Verfahren verwendet werden.

30 Die erfindungsgemäßen Polypeptide der Sequenzen Seq. ID No 39-63, 68-71 können auch als Tool zum Auffinden von Wirkstoffen gegen den Pankreastumor verwendet werden, was ebenfalls Gegenstand der vorliegenden Erfindung ist.

35 Ebenfalls Gegenstand der vorliegenden Erfindung ist die Verwendung der Nukleinsäure-Sequenzen gemäß den Sequenzen Seq. ID No 2-37, 67 zur Expression von Polypeptiden, die als Tools zum Auffinden von Wirkstoffen gegen den Pankreastumor verwendet werden können.

40 Die Erfindung betrifft auch die Verwendung der gefundenen Polypeptid-Teilsequenzen Seq. ID No 39-63, 68-71 als Arzneimittel in der Gentherapie zur Behandlung gegen den Pankreastumor, bzw. zur Herstellung eines Arzneimittels zur Behandlung gegen den Pankreastumor.

45 Die Erfindung betrifft auch Arzneimittel, die mindestens eine Polypeptid-Teilsequenz Seq. ID No 39-63, 68-71 enthalten.

Die gefundenen erfindungsgemäßen Nukleinsäure-Sequenzen können auch genomische oder mRNA-Sequenzen sein.

50 Die Erfindung betrifft auch genomische Gene, ihre Exon- und Intronstruktur und deren Spleißvarianten, erhältlich aus den cDNAs der Sequenzen Seq. ID No 2-37, 67, sowie deren Verwendung zusammen mit geeigneten regulativen Elementen, wie geeigneten Promotoren und/ oder Enhancern.

Mit den erfindungsgermäßen Nukleinsäuren (cDNA-Sequenzen) Seq. ID No 2-37, 67 werden genomische BAC-, PAC- und Cosmid-Bibliotheken gescreent und über komplementäre Basenpaarung (Hybridisierung) spezifisch humane Klone isoliert. Die so isolierten BAC-, PAC- und Cosmid-Klone werden mit Hilfe der Fluoreszenz-in-situ-  
5 Hybridisation auf Metaphasenchromosomen hybridisiert und entsprechende Chromosomenabschnitte identifiziert, auf denen die entsprechenden genomischen Gene liegen. BAC-, PAC- und Cosmid-Klone werden sequenziert, um die entsprechenden genomischen Gene in ihrer vollständigen Struktur (Promotoren, Enhancer, Silencer, Exons und Introns) aufzuklären. BAC-, PAC- und Cosmid-Klone können als eigenständige Moleküle für den Gentransfer eingesetzt werden (s. Fig. 5).  
10 Die Erfindung betrifft auch BAC-, PAC- und Cosmid-Klone, enthaltend funktionelle Gene und ihre chromosomale Lokalisation, entsprechend den Sequenzen Seq. ID No 2-37, 67, zur Verwendung als Vehikel zum Gentransfer.

**Bedeutungen von Fachbegriffen und Abkürzungen**

5 Nukleinsäuren= Unter Nukleinsäuren sind in der voliegenden Erfindung zu verstehen: mRNA, partielle cDNA, vollängen cDNA und genomische Gene (Chromosomen).

10 ORF = Open Reading Frame, eine definierte Abfolge von Aminosäuren, die von der cDNA-Sequenz abgeleitet werden kann.

15 Contig = eine Menge von DNA-Sequenzen, die aufgrund sehr großer Ähnlichkeiten zu einer Sequenz zusammengefaßt werden können (Consensus)

Singleton= ein Contig, der nur eine Sequenz enthält

20 Modul = Domäne eines Proteins mit einer definierten Sequenz, die eine strukturelle Einheit darstellt und in unterschiedlichen Proteinen vorkommt

25 N = wahlweise das Nukleotid A, T, G oder C

X = wahlweise eine der 20 natürlich vorkommenden Aminosäuren

**Erklärung zu den Alignmentparametern**

30 minimal initial match= minimaler anfänglicher Identitätsbereich

maximum pads per read= maximale Anzahl von Insertionen

maximum percent mismatch= maximale Abweichung in %

**Erklärung der Abbildungen**

35 Fig. 1 zeigt die systematische Gen-Suche in der Incyte LifeSeq Datenbank.

40 Fig. 2a zeigt das Prinzip der EST-Assemblierung

Fig. 2b1-2b4 zeigt das gesamte Prinzip der EST-Assemblierung

45 Fig. 3 zeigt die in silico Subtraktion der Genexpression in verschiedenen Geweben

Fig. 4a zeigt die Bestimmung der gewebsspezifischen Expression über elektronischen Northern.

50 Fig. 4b zeigt den elektronischen Northern

Fig. 5 zeigt die Isolierung von genomischen BAC- und PAC-Klonen.

Die nachfolgenden Beispiele erläutern die Herstellung der erfindungsgemäßen Nukleinsäure-Sequenzen, ohne die Erfahrung auf diese Beispiele und Nukleinsäure-Sequenzen zu beschränken.

5

### Beispiel 1

#### Suche nach Tumor-bezogenen Kandidatengenen

10 Zuerst wurden sämtliche ESTs des entsprechenden Gewebes aus der LifeSeq-Datenbank (vom Oktober 1997) extrahiert. Diese wurden dann mittels des Programms GAP4 des Staden-Pakets mit den Parametern 0% mismatch, 8 pads per read und einem minimalen match von 20 assembliert. Die nicht in die GAP4-Datenbank aufgenommenen Sequenzen (Fails) wurden erst bei 1% mismatch und dann nochmals bei 2% mismatch mit der Datenbank assembliert. Aus den Contigs der Datenbank, die aus mehr als einer Sequenz bestanden, wurden Consensussequenzen errechnet. Die Singletons der Datenbank, die nur aus einer Sequenz bestanden, wurden mit den nicht in die GAP4-Datenbank aufgenommenen Sequenzen bei 2% mismatch erneut assembliert. Wiederum wurden für die Contigs die Consensussequenzen ermittelt. Alle übrigen ESTs wurden bei 4% mismatch erneut assembliert. Die Consensussequenzen wurden abermals extrahiert und mit den vorherigen Consensussequenzen sowie den Singletons und den nicht in die Datenbank aufgenommenen Sequenzen abschließend bei 4% mismatch assembliert. 15 20 25 30 35 40 45 50 Die Consensussequenzen wurden gebildet und mit den Singletons und Fails als Ausgangsbasis für die Gewebsvergleiche verwendet. Durch diese Prozedur konnte sichergestellt werden, daß unter den verwendeten Parametern sämtliche Sequenzen von einander unabhängige Genbereiche darstellten.

Fig. 2b1-2b4 veranschaulicht die Verlängerung der Pankreasnormalgewebe ESTs.

Die so assemblierten Sequenzen der jeweiligen Gewebe wurden anschließend mittels des gleichen Programms miteinander verglichen (Fig. 3). Hierzu wurden erst alle Sequenzen des ersten Gewebes in die Datenbank eingegeben. (Daher war es wichtig, daß diese voneinander unabhängig waren.)

Dann wurden alle Sequenzen des zweiten Gewebes mit allen des ersten verglichen. Das Ergebnis waren Sequenzen, die für das erste bzw. das zweite Gewebe spezifisch waren, sowie welche, die in beiden vorkamen. Bei Letzteren wurde das Verhältnis der Häufigkeit des Vorkommens in den jeweiligen Geweben ausgewertet. Sämtliche, die Auswertung der assemblierten Sequenzen betreffenden Programme, wurden selbst entwickelt.

Alle Sequenzen, die mehr als viermal in jeweils einem der verglichenen Gewebe vorkamen, sowie alle, die mindestens fünfmal so häufig in einem der beiden Gewebe vorkamen wurden weiter untersucht. Diese Sequenzen wurden einem elektronischen Northern (s. Beispiel 2.1) unterzogen, wodurch die Verteilung in sämtlichen Tumor- und Normal-Geweben untersucht wurde (s. Fig. 4a und Fig. 4b). Die relevanten Kandidaten wurden dann mit Hilfe sämtlicher Incyte ESTs und allen ESTs öffentlicher Datenbanken verlängert (s. Beispiel 3). Anschließend wurden die Sequenzen und ihre Übersetzung in mögliche Proteine mit allen Nukleotid- und Proteindatenbanken verglichen, sowie auf mögliche, für Proteine kodierende Regionen untersucht.

**Beispiel 2****Algorithmus zur Identifikation und Verlängerung von partiellen cDNA-Sequenzen mit verändertem Expressionsmuster**

5 Im folgenden soll ein Algorithmus zur Auffindung über- oder unterexprimierter Gene erläutert werden. Die einzelnen Schritte sind der besseren Übersicht halber auch in einem Flußdiagramm zusammengefaßt (s. Fig. 4b).

10

**2.1 Elektronischer Northern-Blot**

Zu einer partiellen DNA-Sequenz S, z. B. einem einzelnen EST oder einem Contig von ESTs, werden mittels eines Standardprogramms zur Homologiesuche, z. B. BLAST (Altschul, S. F., Gish W., Miller, W., Myers, E. W. und Lipman, D. J. (1990) *J. Mol. Biol.*, **215**, 403-410), BLAST2 (Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. und Lipman, D. J. (1997) *Nucleic Acids Research* **25** 3389-3402) oder FASTA (Pearson, W. R. und Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85** 2444-2448), die homologen Sequenzen in verschiedenen nach Geweben geordneten (privaten oder öffentlichen) EST-Bibliotheken bestimmt. Die dadurch ermittelten (relativen oder absoluten) Gewebe-spezifischen Vorkommenshäufigkeiten dieser Partial-Sequenz S werden als elektronischer Northern-Blot bezeichnet.

25

**2.1.1**

Analog der unter 2.1 beschriebenen Verfahrensweise wurde die Sequenz Seq. ID No. 2 gefunden, die 9 x häufiger im Pankreasnormalgewebe als im Tumor vorkommt.

Das Ergebnis ist wie folgt:

**Elektronischer Northern für SEQ. ID. NO: 2**

		NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhältnisse N/T    T/N
	Blase	0.0000	0.0026	0.0000 undef
	Brust	0.0000	0.0000	undef undef
	Duenndarm	0.1502	0.0165	9.0831 0.1101
	Eierstock	0.0030	0.0026	1.1513 0.8686
	Endokrines Gewebe	0.0000	0.0000	undef undef
	Gastrointestinal	0.0939	0.0093	10.1472 0.0985
	Gehirn	0.0000	0.0010	0.0000 undef
35	Haematopoetisch	0.0000	0.0000	undef undef
	Haut	0.0000	0.0000	undef undef
	Hepatisch	0.0095	0.0194	0.4902 2.0400
	Herz	0.0000	0.0000	undef undef
	Hoden	0.0000	0.0000	undef undef
40	Lunge	0.0000	0.0000	undef undef
	Magen-Speiseroehre	0.0000	0.1533	0.0000 undef
	Muskel-Skelett	0.0000	0.0000	undef undef
	Niere	0.0000	0.0068	0.0000 undef
45	Pankreas	2.7490	0.3037	9.0506 0.1105
	Penis	0.0000	0.0000	undef undef
50				
55				

	Prostata	0.0000	0.0000	undef	undef
	Uterus_Endometrium	0.0000	0.0000	undef	undef
	Uterus_Myometrium	0.0000	0.0000	undef	undef
5	Uterus_allgemein	0.0000	0.0000	undef	undef
	Brust-Hyperplasie	0.0000			
	Prostata-Hyperplasie	0.0000			
	Samenblase	0.0000			
	Sinnesorgane	0.0000			
10	Weisse_Blutkoerperchen	0.0000			
	Zervix	0.0000			

	FOETUS	
	%Haeufigkeit	
15	Entwicklung	0.0000
	Gastrointenstinal	0.0611
	Gehirn	0.0000
	Haematopoetisch	0.0039
	Haut	0.0000
20	Hepatisch	0.0000
	Herz-Blutgefaesse	0.0036
	Lunge	0.0000
	Nebenniere	0.0000
	Niere	0.0000
25	Placenta	0.0000
	Prostata	0.0000
	Sinnesorgane	0.0000

30	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN	
	%Haeufigkeit	
	Brust	0.0000
	Eierstock_n	0.0000
	Eierstock_t	0.0000
35	Endokrines_Gewebe	0.0000
	Foetal	0.0000
	Gastrointestinal	0.1953
	Haematopoetisch	0.0000
	Haut-Muskel	0.0000
40	Hoden	0.0000
	Lunge	0.0000
	Nerven	0.0000
	Prostata	0.0000
	Sinnesorgane	0.0000
45	Uterus_n	0.0000

In analoger Verfahrensweise wurden auch folgende Northerns gefunden:

#### Elektronischer Northern für SEQ. ID. NO: 4

	NORMAL	TUMOR	Verhaeltnisse		
	%Haeufigkeit	%Haeufigkeit	N/T	T/N	
50	Blase	0.0000	0.0000	undef	undef
	Brust	0.0000	0.0000	undef	undef
	Duenndarm	0.0031	0.0000	undef	0.0000
55	Eierstock	0.0000	0.0000	undef	undef
	Endokrines_Gewebe	0.0000	0.0000	undef	undef
	Gastrointestinal	0.0038	0.0000	undef	0.0000
	Gehirn	0.0000	0.0000	undef	undef
60	Haematopoetisch	0.0000	0.0000	undef	undef
	Haut	0.0000	0.0000	undef	undef
	Hepatisch	0.0000	0.0000	undef	undef
	Herz	0.0000	0.0000	undef	undef
	Hoden	0.0000	0.0000	undef	undef
	Lunge	0.0000	0.0000	undef	undef
65	Magen-Speiseroehre	0.0000	0.0000	undef	undef
	Muskel-Skelett	0.0000	0.0000	undef	undef
	Niere	0.0027	0.0068	0.3965	2.5219

	Pankreas	0.1355	0.0166	8.1767	0.1223
	Penis	0.0000	0.0000	undef	undef
	Prostata	0.0000	0.0021	0.0000	undef
5	Uterus_Endometrium	0.0000	0.0000	undef	undef
	Uterus_Myometrium	0.0000	0.0000	undef	undef
	Uterus_allgemein	0.0000	0.0000	undef	undef
	Brust-Hyperplasie	0.0000			
	Prostata-Hyperplasie	0.0000			
10	Samenblase	0.0000			
	Sinnesorgane	0.0000			
	Weisse_Blutkoerperchen	0.0000			
	Zervix	0.0000			

15	FOETUS	
	%Haeufigkeit	
	Entwicklung	0.0000
	Gastrointestinal	0.0028
	Gehirn	0.0000
20	Haematopoetisch	0.0000
	Haut	0.0000
	Hepatisch	0.0000
	Herz-Blutgefaesse	0.0000
	Lunge	0.0000
25	Nebenniere	0.0000
	Niere	0.0000
	Placenta	0.0000
	Prostata	0.0000
	Sinnesorgane	0.0000

30	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN	
	%Haeufigkeit	
	Brust	0.0000
35	Eierstock_n	0.0000
	Eierstock_t	0.0000
	Endokrines_Gewebe	0.0000
	Foetal	0.0000
	Gastrointestinal	0.0000
40	Haematopoetisch	0.0000
	Haut-Muskel	0.0000
	Hoden	0.0000
	Lunge	0.0000
	Nerven	0.0000
45	Prostata	0.0000
	Sinnesorgane	0.0000
	Uterus_n	0.0000

## Elektronischer Northern für SEQ. ID. NO: 6

	NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse
			N/T    T/N
5	Blase 0.0000	0.0026	0.0000 undef
	Brust 0.0000	0.0000	undef undef
	Duenndarm 0.0031	0.0000	undef 0.0000
	Eierstock 0.0000	0.0000	undef undef
10	Endokrines_Gewebe 0.0000	0.0000	undef undef
	Gastrointestinal 0.0000	0.0000	undef undef
	Gehirn 0.0000	0.0000	undef undef
	Haematopoetisch 0.0000	0.0000	undef undef
	Haut 0.0000	0.0000	undef undef
	Hepatisch 0.0000	0.0000	undef undef
15	Herz 0.0000	0.0000	undef undef
	Hoden 0.0000	0.0000	undef undef
	Lunge 0.0062	0.0000	undef 0.0000
	Magen-Speiseroehre 0.0000	0.0000	undef undef
	Muskel-Skelett 0.0000	0.0000	undef undef
20	Niere 0.0000	0.0000	undef undef
	Pankreas 0.8078	0.0166	48.7611 0.0205
	Penis 0.0000	0.0000	undef undef
	Prostata 0.0000	0.0000	undef undef
25	Uterus_Endometrium 0.0000	0.0000	undef undef
	Uterus_Myometrium 0.0000	0.0000	undef undef
	Uterus_allgemein 0.0000	0.0000	undef undef
	Brust-Hyperplasie 0.0000		
	Prostata-Hyperplasie 0.0000		
	Samenblase 0.0000		
30	Sinnesorgane 0.0000		
	Weisse_Blutkoerperchen 0.0000		
	Zervix 0.0000		

	FOETUS %Haeufigkeit
35	Entwicklung 0.0000
	Gastrointestinal 0.0139
	Gehirn 0.0000
40	Haematopoetisch 0.0000
	Haut 0.0000
	Hepatisch 0.0000
	Herz-Blutgefaesse 0.0000
	Lunge 0.0000
45	Nebenniere 0.0000
	Niere 0.0000
	Placenta 0.0000
	Prostata 0.0000
50	Sinnesorgane 0.0000

	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN %Haeufigkeit
55	Brust 0.0000
	Eierstock_n 0.0000
	Eierstock_t 0.0000
	Endokrines_Gewebe 0.0000
	Foetal 0.0000
60	Gastrointestinal 0.0000
	Haematopoetisch 0.0000
	Haut-Muskel 0.0000
	Hoden 0.0000
	Lunge 0.0000
	Nerven 0.0000
65	Prostata 0.0000
	Sinnesorgane 0.0000
	Uterus_n 0.0000

## Elektronischer Northern für SEQ. ID. NO: 7

		NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse N/T    T/N
5	Blase	0.0000	0.0000	undef undef
	Brust	0.0000	0.0000	undef undef
	Duenndarm	0.0031	0.0000	undef 0.0000
	Eierstock	0.0000	0.0000	undef undef
10	Endokrines_Gewebe	0.0000	0.3310	0.0000 undef
	Gastrointestinal	0.0019	0.0000	undef 0.0000
	Gehirn	0.0089	0.0051	1.7279 0.5787
	Haematopoetisch	0.0000	0.0000	undef undef
	Haut	0.0000	0.0000	undef undef
15	Hepatisch	0.0000	0.0000	undef undef
	Herz	0.0000	0.0000	undef undef
	Hoden	0.0000	0.0000	undef undef
	Lunge	0.0000	0.0000	undef undef
	Magen-Speiseroehre	0.0000	0.0000	undef undef
	Muskel-Skelett	0.0000	0.0000	undef undef
20	Niere	0.0000	0.0000	undef undef
	Pankreas	0.0991	0.0055	17.9489 0.0557
	Penis	0.0000	0.0000	undef undef
	Prostata	0.0000	0.0000	undef undef
25	Uterus_Endometrium	0.0000	0.0000	undef undef
	Uterus_Myometrium	0.0000	0.0000	undef undef
	Uterus_allgemein	0.0000	0.0000	undef undef
	Brust-Hyperplasie	0.0000		
	Prostata-Hyperplasie	0.0000		
	Samenblase	0.0000		
30	Sinnesorgane	0.0000		
	Weisse_Blutkoerperchen	0.0000		
	Zervix	0.0000		
35		FOETUS %Haeufigkeit		
	Entwicklung	0.0000		
	Gastrointenstinal	0.0389		
	Gehirn	0.0000		
40	Haematopoetisch	0.0000		
	Haut	0.0000		
	Hepatisch	0.0000		
	Herz-Blutgefaesse	0.0000		
	Lunge	0.0000		
45	Nebenniere	0.0000		
	Niere	0.0000		
	Placenta	0.0000		
	Prostata	0.0000		
50	Sinnesorgane	0.0000		
		NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN %Haeufigkeit		
	Brust	0.0000		
55	Eierstock_n	0.0000		
	Eierstock_t	0.0000		
	Endokrines_Gewebe	0.0000		
	Foetal	0.0006		
	Gastrointestinal	0.0000		
60	Haematopoetisch	0.0000		
	Haut-Muskel	0.0000		
	Hoden	0.0000		
	Lunge	0.0000		
	Nerven	0.0110		
65	Prostata	0.0000		
	Sinnesorgane	0.0000		
	Uterus_n	0.0000		

## Elektronischer Northern für SEQ. ID. NO: 9

		NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse N/T    T/N
5	Blase	0.0000	0.0000	undef undef
	Brust	0.0000	0.0000	undef undef
	Duenndarm	0.0000	0.0000	undef undef
	Eierstock	0.0000	0.0000	undef undef
10	Endokrines_Gewebe	0.0000	0.0000	undef undef
	Gastrointestinal	0.0000	0.0000	undef undef
	Gehirn	0.0000	0.0000	undef undef
	Haematopoetisch	0.0000	0.0000	undef undef
	Haut	0.0000	0.0000	undef undef
15	Hepatisch	0.0000	0.0000	undef undef
	Herz	0.0000	0.0000	undef undef
	Hoden	0.0000	0.0000	undef undef
	Lunge	0.0000	0.0000	undef undef
20	Magen-Speiseroehre	0.0000	0.0000	undef undef
	Muskel-Skelett	0.0000	0.0000	undef undef
	Niere	0.0000	0.0000	undef undef
	Pankreas	0.1867	0.0000	undef 0.0000
	Penis	0.0000	0.0000	undef undef
	Prostata	0.0000	0.0000	undef undef
25	Uterus_Endometrium	0.0000	0.0000	undef undef
	Uterus_Myometrium	0.0000	0.0000	undef undef
	Uterus_allgemein	0.0000	0.0000	undef undef
	Brust-Hyperplasie	0.0000		
	Prostata-Hyperplasie	0.0000		
	Samenblase	0.0000		
30	Sinnesorgane	0.0000		
	Weisse_Blutkoerperchen	0.0000		
	Zervix	0.0000		
35	FOETUS %Haeufigkeit			
	Entwicklung	0.0000		
	Gastrointenstinal	0.0000		
	Gehirn	0.0000		
40	Haematopoetisch	0.0000		
	Haut	0.0000		
	Hepatisch	0.0000		
	Herz-Blutgefaesse	0.0000		
	Lunge	0.0000		
45	Nebenniere	0.0000		
	Niere	0.0000		
	Placenta	0.0000		
	Prostata	0.0000		
50	Sinnesorgane	0.0000		
	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN			
	%Haeufigkeit			
55	Brust	0.0000		
	Eierstock_n	0.0000		
	Eierstock_t	0.0000		
	Endokrines_Gewebe	0.0000		
	Foetal	0.0000		
60	Gastrointestinal	0.0244		
	Haematopoetisch	0.0000		
	Haut-Muskel	0.0000		
	Hoden	0.0154		
	Lunge	0.0000		
	Nerven	0.0000		
65	Prostata	0.0000		
	Sinnesorgane	0.0000		
	Uterus_n	0.0000		

## Elektronischer Northern für SEQ. ID. NO: 11

		NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse N/T      T/N
5	Blase	0.0000	0.0128	0.0000 undef
	Brust	0.0000	0.0000	undef undef
	Duenndarm	0.0000	0.0000	undef undef
	Eierstock	0.0060	0.0234	0.2558 3.9088
10	Endokrines_Gewebe	0.0000	0.0000	undef undef
	Gastrointestinal	0.0096	0.0000	undef 0.0000
	Gehirn	0.0000	0.0000	undef undef
	Haematopoetisch	0.0000	0.0000	undef undef
	Haut	0.0000	0.0000	undef undef
15	Hepatisch	0.0190	0.0000	undef 0.0000
	Herz	0.0011	0.0000	undef 0.0000
	Hoden	0.0000	0.0000	undef undef
	Lunge	0.0000	0.0020	0.0000 undef
	Magen-Speiseroehre	0.0097	0.0077	1.2605 0.7933
20	Muskel-Skelett	0.0000	0.0000	undef undef
	Niere	0.0000	0.0000	undef undef
	Pankreas	0.4180	0.0055	75.6843 0.0132
	Penis	0.0000	0.0000	undef undef
	Prostata	0.0000	0.0000	undef undef
25	Uterus_Endometrium	0.0000	0.0000	undef undef
	Uterus_Myometrium	0.0000	0.0000	undef undef
	Uterus_allgemein	0.0000	0.0000	undef undef
	Brust-Hyperplasie	0.0000		
	Prostata-Hyperplasie	0.0000		
	Samenblase	0.0000		
30	Sinnesorgane	0.0000		
	Weisse_Blutkoerperchen	0.0009		
	Zervix	0.0000		

	FOETUS %Haeufigkeit
35	Entwicklung 0.0000
	Gastrointenstinal 0.0944
	Gehirn 0.0000
40	Haematopoetisch 0.0000
	Haut 0.0000
	Hepatisch 0.0000
	Herz-Blutgefaesse 0.0000
	Lunge 0.0000
45	Nebenniere 0.0000
	Niere 0.0000
	Placenta 0.0000
	Prostata 0.0000
50	Sinnesorgane 0.0000

	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN %Haeufigkeit
55	Brust 0.0000
	Eierstock_n 0.0000
	Eierstock_t 0.0000
	Endokrines_Gewebe 0.0000
	Foetal 0.0000
60	Gastrointestinal 0.0000
	Haematopoetisch 0.0000
	Haut-Muskel 0.0000
	Hoden 0.0000
	Lunge 0.0164
	Nerven 0.0000
65	Prostata 0.0000
	Sinnesorgane 0.0000
	Uterus_n 0.0000

## Elektronischer Northern für SEQ. ID. NO: 12

	NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse N/T    T/N
5	Blase 0.0000	0.0051	0.0000 undef
	Brust 0.0000	0.0075	0.0000 undef
	Duenndarm 0.0061	0.0000	undef 0.0000
	Eierstock 0.0000	0.0000	undef undef
10	Endokrines_Gewebe 0.0000	0.0000	undef undef
	Gastricintestinal 0.0000	0.0000	undef undef
	Gehirn 0.0000	0.0041	0.0000 undef
	Haematopoetisch 0.0027	0.0000	undef 0.0000
	Haut 0.0000	0.0000	undef undef
15	Hepatisch 0.0000	0.0000	undef undef
	Herz 0.0032	0.0000	undef 0.0000
	Hoden 0.0000	0.0000	undef undef
	Lunge 0.0000	0.0000	undef undef
	Magen-Speiseroehre 0.0000	0.0000	undef undef
	Muskel-Skelett 0.0000	0.0000	undef undef
20	Niere 0.0000	0.0000	undef undef
	Pankreas 3.8492	0.2706	14.2248 0.0703
	Penis 0.0000	0.0000	undef undef
	Prostata 0.0000	0.0000	undef undef
25	Uterus_Endometrium 0.0000	0.0000	undef undef
	Uterus_Myometrium 0.0000	0.0000	undef undef
	Uterus_allgemein 0.0000	0.0000	undef undef
	Brust-Hyperplasie 0.0000		
	Prostata-Hyperplasie 0.0000		
	Samenblase 0.0000		
30	Sinnesorgane 0.0000		
	Weisse_Blutkoerperchen 0.0009		
	Zervix 0.0000		

	FOETUS %Haeufigkeit
35	Entwicklung 0.0000
	Gastrointenstinal 0.2305
	Gehirn 0.0000
40	Haematopoetisch 0.0039
	Haut 0.0000
	Hepatisch 0.0000
	Herz-Blutgefaesse 0.0000
	Lunge 0.0000
45	Nebenniere 0.0000
	Niere 0.0000
	Placenta 0.0000
	Prostata 0.0000
50	Sinnesorgane 0.0000

	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN %Haeufigkeit	
55	Brust 0.0000	
	Eierstock_n 0.0000	
	Eierstock_t 0.0000	
	Endokrines_Gewebe 0.0000	
	Foetal 0.0000	
60	Gastrointestinal 0.4149	
	Haematopoetisch 0.0000	
	Haut-Muskel 0.0000	
	Hoden 0.0000	
	Lunge 0.0000	
	Nerven 0.0000	
65	Prostata 0.0000	
	Sinnesorgane 0.0000	
	Uterus_n 0.0000	

## Elektronischer Northern für SEQ. ID. NO: 14

	NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse	N/T	T/N
5	Blase 0.0000	0.0000	undef	undef	
	Brust 0.0000	0.0000	undef	undef	
	Duenndarm 0.0000	0.0000	undef	undef	
	Eierstock 0.0000	0.0000	undef	undef	
10	Endokrines_Gewebe 0.0000	0.0000	undef	undef	
	Gastrointestinal 0.0000	0.0000	undef	undef	
	Gehirn 0.0000	0.0000	undef	undef	
	Haematopoetisch 0.0000	0.0000	undef	undef	
	Haut 0.0000	0.0000	undef	undef	
15	Hepatisch 0.0000	0.0000	undef	undef	
	Herz 0.0000	0.0000	undef	undef	
	Hoden 0.0000	0.0000	undef	undef	
	Lunge 0.0000	0.0000	undef	undef	
	Magen-Speiseroehre 0.0000	0.0000	undef	undef	
	Muskel-Skelett 0.0000	0.0000	undef	undef	
20	Niere 0.0000	0.0000	undef	undef	
	Pankreas 0.0182	0.0000	undef	0.0000	
	Penis 0.0000	0.0000	undef	undef	
	Prostata 0.0000	0.0000	undef	undef	
25	Uterus_Endometrium 0.0000	0.0000	undef	undef	
	Uterus_Myometrium 0.0000	0.0000	undef	undef	
	Uterus_allgemein 0.0000	0.0000	undef	undef	
	Brust-Hyperplasie 0.0000				
	Prostata-Hyperplasie 0.0000				
	Samenblase 0.0000				
30	Sinnesorgane 0.0000				
	Weisse_Blutkoerperchen 0.0000				
	Zervix 0.0000				
35	FOETUS %Haeufigkeit				
	Entwicklung 0.0000				
	Gastrointenstinal 0.0028				
	Gehirn 0.0000				
40	Haematopoetisch 0.0000				
	Haut 0.0000				
	Hepatisch 0.0000				
	Herz-Blutgefaesse 0.0000				
	Lunge 0.0000				
45	Nebenniere 0.0000				
	Niere 0.0000				
	Placenta 0.0000				
	Prostata 0.0000				
	Sinnesorgane 0.0000				
50	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN %Haeufigkeit				
	Brust 0.0000				
55	Eierstock_n 0.0000				
	Eierstock_t 0.0000				
	Endokrines_Gewebe 0.0000				
	Foetal 0.0000				
60	Gastrointestinal 0.0000				
	Haematopoetisch 0.0000				
	Haut-Muskel 0.0000				
	Hoden 0.0000				
	Lunge 0.0000				
	Nerven 0.0000				
65	Prostata 0.0000				
	Sinnesorgane 0.0000				
	Uterus_n 0.0000				

## Elektronischer Northern für SEQ. ID. NO: 15

		NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse N/T    T/N
5	Blase	0.0000	0.0000	undef    undef
	Brust	0.0000	0.0038	0.0000    undef
	Duenndarm	0.0061	0.0000	undef    0.0000
	Eierstock	0.0000	0.0000	undef    undef
10	Endokrines_Gewebe	0.0000	0.0000	undef    undef
	Gastrointestinal	0.0000	0.0000	undef    undef
	Gehirn	0.0000	0.0051	0.0000    undef
	Haematopoetisch	0.0040	0.0000	undef    0.0000
	Haut	0.0000	0.0000	undef    undef
15	Hepatisch	0.0000	0.0000	undef    undef
	Herz	0.0011	0.0000	undef    0.0000
	Hoden	0.0000	0.0000	undef    undef
	Lunge	0.0000	0.0000	undef    undef
	Magen-Speiseroehre	0.0000	0.0000	undef    undef
	Muskel-Skelett	0.0000	0.0000	undef    undef
20	Niere	0.0000	0.0000	undef    undef
	Pankreas	2.8861	0.2872	10.0502    0.0995
	Penis	0.0000	0.0000	undef    undef
	Prostata	0.0000	0.0000	undef    undef
25	Uterus_Endometrium	0.0000	0.0000	undef    undef
	Uterus_Myometrium	0.0000	0.0000	undef    undef
	Uterus_allgemein	0.0000	0.0000	undef    undef
	Brust-Hyperplasie	0.0000		
	Prostata-Hyperplasie	0.0000		
	Samenblase	0.0000		
30	Sinnesorgane	0.0000		
	Weisse_Blutkoerperchen	0.0009		
	Zervix	0.0000		

	FOETUS %Haeufigkeit
35	Entwicklung 0.0000
	Gastrointestinal 0.2777
	Gehirn 0.0000
40	Haematopoetisch 0.0039
	Haut 0.0000
	Hepatisch 0.0000
	Herz-Blutgefaesse 0.0000
	Lunge 0.0000
45	Nebenniere 0.0000
	Niere 0.0000
	Placenta 0.0000
	Prostata 0.0000
	Sinnesorgane 0.0000

	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN %Haeufigkeit
50	Brust 0.0000
55	Eierstock_n 0.0000
	Eierstock_t 0.0000
	Endokrines_Gewebe 0.0000
	Foetal 0.0000
60	Gastrointestinal 0.0854
	Haematopoetisch 0.0000
	Haut-Muskel 0.0000
	Hoden 0.0000
	Lunge 0.0000
	Nerven 0.0000
65	Prostata 0.0000
	Sinnesorgane 0.0000
	Uterus_n 0.0000

## Elektronischer Northern für SEQ. ID. NO: 16

		NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse N/T	T/N
5	Blase	0.0000	0.0026	0.0000	undef
	Brust	0.0000	0.0000	undef	undef
	Duenndarm	0.0061	0.0000	undef	0.0000
	Eierstock	0.0000	0.0000	undef	undef
10	Endokrines_Gewebe	0.0000	0.0000	undef	undef
	Gastrointestinal	0.0000	0.0000	undef	undef
	Gehirn	0.0000	0.0051	0.0000	undef
	Haematopoetisch	0.0000	0.0000	undef	undef
	Haut	0.0000	0.0000	undef	undef
15	Hepatisch	0.0000	0.0000	undef	undef
	Herz	0.0011	0.0000	undef	0.0000
	Hoden	0.0000	0.0000	undef	undef
	Lunge	0.0010	0.0000	undef	0.0000
	Magen-Speiseroehre	0.0000	0.0000	undef	undef
	Muskel-Skelett	0.0000	0.0000	undef	undef
20	Niere	0.0000	0.0068	0.0000	undef
	Pankreas	1.5232	0.0110	137.9070	0.0073
	Penis	0.0000	0.0000	undef	undef
	Prostata	0.0000	0.0000	undef	undef
25	Uterus_Endometrium	0.0000	0.0000	undef	undef
	Uterus_Myometrium	0.0000	0.0000	undef	undef
	Uterus_allgemein	0.0000	0.0000	undef	undef
	Brust-Hyperplasie	0.0000			
	Prostata-Hyperplasie	0.0000			
	Samenblase	0.0000			
30	Sinnesorgane	0.0000			
	Weisse_Blutkoerperchen	0.0000			
	Zervix	0.0000			

		FOETUS %Haeufigkeit
35	Entwicklung	0.0000
	Gastrointestinal	0.6386
	Gehirn	0.0000
40	Haematopoetisch	0.0197
	Haut	0.0000
	Hepatisch	0.0000
	Herz-Blutgefaesse	0.0000
	Lunge	0.0000
45	Nebenniere	0.0000
	Niere	0.0000
	Placenta	0.0000
	Prostata	0.0000
50	Sinnesorgane	0.0000

		NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN %Haeufigkeit
55	Brust	0.0000
	Eierstock_n	0.0000
	Eierstock_t	0.0000
	Endokrines_Gewebe	0.0000
	Foetal	0.0000
60	Gastrointestinal	0.0000
	Haematopoetisch	0.0000
	Haut-Muskel	0.0000
	Hoden	0.0000
	Lunge	0.0000
	Nerven	0.0000
65	Prostata	0.0000
	Sinnesorgane	0.0000
	Uterus_n	0.0000

## Elektronischer Northern für SEQ. ID. NO: 17

		NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse N/T	T/N	
5	Blase	0.0000	0.0051	0.0000	undef	
	Brust	0.0000	0.0056	0.0000	undef	
	Duenndarm	0.0000	0.0000	undef	undef	
	Eierstock	0.0000	0.0026	0.0000	undef	
10	Endokrines_Gewebe	0.0000	0.0000	undef	undef	
	Gastrointestinal	0.0000	0.0000	undef	undef	
	Gehirn	0.0000	0.0041	0.0000	undef	
	Haematopoetisch	0.0027	0.0000	undef	0.0000	
	Haut	0.0000	0.0000	undef	undef	
15	Hepatisch	0.0000	0.0000	undef	undef	
	Herz	0.0000	0.0000	undef	undef	
	Hoden	0.0000	0.0000	undef	undef	
	Lunge	0.0010	0.0000	undef	0.0000	
	Magen-Speiseroehre	0.0000	0.0000	undef	undef	
	Muskel-Skelett	0.0000	0.0000	undef	undef	
20	Niere	0.0000	0.0000	undef	undef	
	Pankreas	0.9218	0.0387	23.8463	0.0419	
	Penis	0.0000	0.0000	undef	undef	
	Prostata	0.0000	0.0000	undef	undef	
25	Uterus_Endometrium	0.0000	0.0000	undef	undef	
	Uterus_Myometrium	0.0000	0.0000	undef	undef	
	Uterus_allgemein	0.0000	0.0000	undef	undef	
	Brust-Hyperplasie	0.0000				
	Prostata-Hyperplasie	0.0000				
	Samenblase	0.0000				
30	Sinnesorgane	0.0000				
	Weisse_Blutkoerperchen	0.0000				
	Zervix	0.0000				
35			FOETUS			
			%Haeufigkeit			
			Entwicklung	0.0000		
			Gastrointenstinal	0.1111		
			Gehirn	0.0000		
40			Haematopoetisch	0.0157		
			Haut	0.0000		
			Hepatisch	0.0000		
			Herz-Blutgefaesse	0.0000		
			Lunge	0.0000		
45			Nebenniere	0.0000		
			Niere	0.0000		
			Placenta	0.0000		
			Prostata	0.0000		
50			Sinnesorgane	0.0000		
					NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN	
					%Haeufigkeit	
					Brust	0.0000
55					Eierstock_n	0.0000
					Eierstock_t	0.0000
					Endokrines_Gewebe	0.0000
					Foetal	0.0000
					Gastrointestinal	0.0000
60					Haematopoetisch	0.0000
					Haut-Muskel	0.0000
					Hoden	0.0000
					Lunge	0.0000
					Nerven	0.0000
65					Prostata	0.0000
					Sinnesorgane	0.0000
					Uterus_n	0.0000

## Elektronischer Northern für SEQ. ID. NO: 18

	NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse	
			N/T	T/N
5	Blase 0.0000	0.0026	0.0000	undef
	Brust 0.0000	0.0000	undef	undef
	Duenndarm 0.0000	0.0000	undef	undef
	Eierstock 0.0030	0.0000	undef	0.0000
10	Endokrines_Gewebe 0.0000	0.0000	undef	undef
	Gastrointestinal 0.0000	0.0000	undef	undef
	Gehirn 0.0000	0.0010	0.0000	undef
	Haematopoetisch 0.0000	0.0000	undef	undef
	Haut 0.0000	0.0000	undef	undef
15	Hepatisch 0.0000	0.0000	undef	undef
	Herz 0.0000	0.0000	undef	undef
	Hoden 0.0000	0.0000	undef	undef
	Lunge 0.0000	0.0000	undef	undef
	Magen-Speiseroehre 0.0000	0.0000	undef	undef
	Muskel-Skelett 0.0000	0.0000	undef	undef
20	Niere 0.0054	0.0000	undef	0.0000
	Pankreas 0.4659	0.0221	21.0899	0.0474
	Penis 0.0000	0.0000	undef	undef
	Prostata 0.0000	0.0000	undef	undef
25	Uterus_Endometrium 0.0000	0.0000	undef	undef
	Uterus_Myometrium 0.0000	0.0000	undef	undef
	Uterus_allgemein 0.0000	0.0000	undef	undef
	Brust-Hyperplasie 0.0000			
	Prostata-Hyperplasie 0.0030			
	Samenblase 0.0000			
30	Sinnesorgane 0.0000			
	Weisse_Blutkoerperchen 0.0000			
	Zervix 0.0000			
35	FOETUS %Haeufigkeit			
	Entwicklung 0.0000			
	Gastrointestinal 0.0305			
	Gehirn 0.0000			
40	Haematopoetisch 0.0039			
	Haut 0.0000			
	Hepatisch 0.0000			
	Herz-Blutgefaesse 0.0000			
	Lunge 0.0000			
45	Nebenniere 0.0000			
	Niere 0.0000			
	Placenta 0.0000			
	Prostata 0.0000			
50	Sinnesorgane 0.0000			
	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN %Haeufigkeit			
	Brust 0.0000			
55	Eierstock_n 0.0000			
	Eierstock_t 0.0000			
	Endokrines_Gewebe 0.0000			
	Foetal 0.0000			
60	Gastrointestinal 0.0000			
	Haematopoetisch 0.0000			
	Haut-Muskel 0.0000			
	Hoden 0.0000			
	Lunge 0.0000			
	Nerven 0.0000			
65	Prostata 0.0000			
	Sinnesorgane 0.0000			
	Uterus_n 0.0000			

## Elektronischer Northern für SEQ. ID. NO: 19

	NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse N/T    T/N
5	Blase 0.0000	0.0000	undef undef
	Brust 0.0000	0.0000	undef undef
	Duenndarm 0.0000	0.0000	undef undef
	Eierstock 0.0000	0.0000	undef undef
10	Endokrines_Gewebe 0.0000	0.0000	undef undef
	Gastrointestinal 0.0000	0.0000	undef undef
	Gehirn 0.0000	0.0000	undef undef
	Haematopoetisch 0.0000	0.0000	undef undef
	Haut 0.0000	0.0000	undef undef
	Hepatisch 0.0000	0.0000	undef undef
15	Herz 0.0000	0.0000	undef undef
	Hoden 0.0000	0.0000	undef undef
	Lunge 0.0000	0.0000	undef undef
	Magen-Speiseroehre 0.0000	0.0000	undef undef
	Muskel-Skelett 0.0000	0.0000	undef undef
20	Niere 0.0000	0.0000	undef undef
	Pankreas 0.0314	0.0000	undef 0.0000
	Penis 0.0000	0.0000	undef undef
	Prostata 0.0000	0.0000	undef undef
25	Uterus_Endometrium 0.0000	0.0000	undef undef
	Uterus_Myometrium 0.0000	0.0000	undef undef
	Uterus_allgemein 0.0000	0.0000	undef undef
	Brust-Hyperplasie 0.0000		
	Prostata-Hyperplasie 0.0000		
	Samenblase 0.0000		
30	Sinnesorgane 0.0000		
	Weisse_Blutkoerperchen 0.0000		
	Zervix 0.0000		

	FOETUS %Haeufigkeit
35	Entwicklung 0.0000
	Gastrointenstinal 0.0000
	Gehirn 0.0000
40	Haematopoetisch 0.0000
	Haut 0.0000
	Hepatisch 0.0000
	Herz-Blutgefaesse 0.0000
	Lunge 0.0000
45	Nebenniere 0.0000
	Niere 0.0000
	Placenta 0.0000
	Prostata 0.0000
	Sinnesorgane 0.0000
50	

	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN %Haeufigkeit
55	Brust 0.0000
	Eierstock_n 0.0000
	Eierstock_t 0.0000
	Endokrines_Gewebe 0.0000
	Foetal 0.0000
60	Gastrointestinal 0.0366
	Haematopoetisch 0.0000
	Haut-Muskel 0.0000
	Hoden 0.0000
	Lunge 0.0000
	Nerven 0.0000
65	Prostata 0.0000
	Sinnesorgane 0.0000
	Uterus_n 0.0000

## Elektronischer Northern für SEQ. ID. NO: 21

	NORMAL		TUMOR		Verhaeltnisse	
	%Haeufigkeit	%Haeufigkeit	%Haeufigkeit	%Haeufigkeit	N/T	T/N
5	Blase	0.0000	0.0026	0.0000	undef	
	Brust	0.0000	0.0000	undef	undef	
	Duenndarm	0.0000	0.0000	undef	undef	
	Eierstock	0.0000	0.0000	undef	undef	
10	Endokrines_Gewebe	0.0000	0.0000	undef	undef	
	Gastrointestinal	0.0000	0.0000	undef	undef	
	Gehirn	0.0000	0.0000	undef	undef	
	Haematopoetisch	0.0000	0.0000	undef	undef	
15	Haut	0.0000	0.0000	undef	undef	
	Hepatisch	0.0000	0.0000	undef	undef	
	Herz	0.0000	0.0000	undef	undef	
	Hoden	0.0000	0.0000	undef	undef	
20	Lunge	0.0000	0.0000	undef	undef	
	Magen-Speiseroehre	0.0000	0.0000	undef	undef	
	Muskel-Skelett	0.0000	0.0000	undef	undef	
	Niere	0.0000	0.0000	undef	undef	
25	Pankreas	0.1586	0.0000	undef	0.0000	
	Penis	0.0000	0.0000	undef	undef	
	Prostata	0.0000	0.0000	undef	undef	
	Uterus_Endometrium	0.0000	0.0000	undef	undef	
30	Uterus_Myometrium	0.0000	0.0000	undef	undef	
	Uterus_allgemein	0.0000	0.0000	undef	undef	
	Brust-Hyperplasie	0.0000				
	Prostata-Hyperplasie	0.0000				
35	Samenblase	0.0000				
	Sinnesorgane	0.0000				
	Weisse_Blutkoerperchen	0.0000				
	Zervix	0.0000				
40	FOETUS					
	%Haeufigkeit					
	Entwicklung	0.0000				
	Gastrointestinal	0.0000				
45	Gehirn	0.0000				
	Haematopoetisch	0.0000				
	Haut	0.0000				
	Hepatisch	0.0000				
50	Herz-Blutgefaesse	0.0000				
	Lunge	0.0000				
	Nebenniere	0.0000				
	Niere	0.0000				
55	Placenta	0.0000				
	Prostata	0.0000				
	Sinnesorgane	0.0000				
	60	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN				
%Haeufigkeit						
Brust		0.0000				
Eierstock_n		0.0000				
65	Eierstock_t	0.0000				
	Endokrines_Gewebe	0.0000				
	Foetal	0.0000				
	Gastrointestinal	0.0000				
70	Haematopoetisch	0.0057				
	Haut-Muskel	0.0000				
	Hoden	0.0000				
	Lunge	0.0000				
75	Nerven	0.0000				
	Prostata	0.0000				
	Sinnesorgane	0.0000				
	Uterus_n	0.0000				

## Elektronischer Northern für SEQ. ID. NO: 24

	NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse
			N/T      T/N
5	Blase 0.0000	0.0000	undef    undef
	Brust 0.0000	0.0000	undef    undef
	Duenndarm 0.0000	0.0000	undef    undef
	Eierstock 0.0000	0.0000	undef    undef
10	Endokrines_Gewebe 0.0000	0.0050	0.0000    undef
	Gastricintestinal 0.0000	0.0000	undef    undef
	Gehirn 0.0007	0.0000	undef    0.0000
	Haematopoetisch 0.0000	0.0000	undef    undef
	Haut 0.0000	0.0000	undef    undef
15	Hepatisch 0.0000	0.0000	undef    undef
	Herz 0.0000	0.0000	undef    undef
	Hoden 0.0000	0.0000	undef    undef
	Lunge 0.0000	0.0000	undef    undef
	Magen-Speiseroehre 0.0000	0.0000	undef    undef
20	Muskel-Skelett 0.0000	0.0000	undef    undef
	Niere 0.0000	0.0000	undef    undef
	Pankreas 0.0248	0.0000	undef    0.0000
	Penis 0.0000	0.0000	undef    undef
	Prostata 0.0000	0.0000	undef    undef
25	Uterus_Endometrium 0.0000	0.0000	undef    undef
	Uterus_Myometrium 0.0000	0.0000	undef    undef
	Uterus_allgemein 0.0000	0.0000	undef    undef
	Brust-Hyperplasie 0.0000		
	Prostata-Hyperplasie 0.0030		
	Samenblase 0.0000		
30	Sinnesorgane 0.0000		
	Weisse_Blutkoerperchen 0.0000		
	Zervix 0.0000		

	FOETUS %Haeufigkeit
35	Entwicklung 0.0000
	Gastrointenstinal 0.0056
	Gehirn 0.0000
40	Haematopoetisch 0.0000
	Haut 0.0000
	Hepatisch 0.0000
	Herz-Blutgefaesse 0.0000
	Lunge 0.0000
45	Nebenniere 0.0000
	Niere 0.0000
	Placenta 0.0000
	Prostata 0.0000
	Sinnesorgane 0.0000

	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN %Haeufigkeit	
50	Brust 0.0000	
55	Eierstock_n 0.0000	
	Eierstock_t 0.0000	
	Endokrines_Gewebe 0.0000	
	Foetal 0.0000	
60	Gastrointestinal 0.0122	
	Haematopoetisch 0.0000	
	Haut-Muskel 0.0000	
	Hoden 0.0000	
	Lunge 0.0000	
	Nerven 0.0010	
65	Prostata 0.0000	
	Sinnesorgane 0.0000	
	Uterus_n 0.0000	

## Elektronischer Northern für SEQ. ID. NO: 25

	NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse N/T      T/N
5	Blase 0.0000	0.0000	undef undef
	Brust 0.0000	0.0019	0.0000 undef
	Duenndarm 0.0000	0.0000	undef undef
	Eierstock 0.0000	0.0000	undef undef
10	Endokrines_Gewebe 0.0000	0.0000	undef undef
	Gastrointestinal 0.0000	0.0000	undef undef
	Gehirn 0.0000	0.0000	undef undef
	Haematopoetisch 0.0000	0.0000	undef undef
	Haut 0.0000	0.0000	undef undef
15	Hepatisch 0.0000	0.0000	undef undef
	Herz 0.0000	0.0000	undef undef
	Hoden 0.0000	0.0000	undef undef
	Lunge 0.0000	0.0000	undef undef
	Magen-Speiseroehre 0.0000	0.0000	undef undef
	Muskel-Skelett 0.0000	0.0000	undef undef
20	Niere 0.0000	0.0000	undef undef
	Pankreas 0.2263	0.0055	40.9832      0.0244
	Penis 0.0000	0.0000	undef undef
	Prostata 0.0000	0.0000	undef undef
25	Uterus_Endometrium 0.0000	0.0000	undef undef
	Uterus_Myometrium 0.0000	0.0000	undef undef
	Uterus_allgemein 0.0000	0.0000	undef undef
	Brust-Hyperplasie 0.0000		
	Prostata-Hyperplasie 0.0000		
	Samenblase 0.0000		
30	Sinnesorgane 0.0000		
	Weisse_Blutkoerperchen 0.0000		
	Zervix 0.0000		

	FOETUS %Haeufigkeit
35	Entwicklung 0.0000
	Gastrointenstinal 0.0528
	Gehirn 0.0000
40	Haematopoetisch 0.0000
	Haut 0.0000
	Hepatisch 0.0000
	Herz-Blutgefaesse 0.0000
	Lunge 0.0000
45	Nebenniere 0.0000
	Niere 0.0000
	Placenta 0.0000
	Prostata 0.0000
50	Sinnesorgane 0.0000

	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN %Haeufigkeit
55	Brust 0.0000
	Eierstock_n 0.0000
	Eierstock_t 0.0000
	Endokrines_Gewebe 0.0000
	Foetal 0.0000
60	Gastrointestinal 0.0000
	Haematopoetisch 0.0000
	Haut-Muskel 0.0000
	Hoden 0.0000
	Lunge 0.0000
	Nerven 0.0000
65	Prostata 0.0000
	Sinnesorgane 0.0000
	Uterus_n 0.0000

## Elektronischer Northern für SEQ. ID. NO: 27

	NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse	
			N/T	T/N
5	Blase 0.0000	0.0000	undef	undef
	Brust 0.0000	0.0000	undef	undef
	Duenndarm 0.0000	0.0000	undef	undef
	Eierstock 0.0000	0.0000	undef	undef
10	Endokrines_Gewebe 0.0000	0.0000	undef	undef
	Gastrointestinal 0.0000	0.0000	undef	undef
	Gehirn 0.0000	0.0000	undef	undef
	Haematopoetisch 0.0000	0.0000	undef	undef
	Haut 0.0000	0.0000	undef	undef
15	Hepatisch 0.0000	0.0000	undef	undef
	Herz 0.0000	0.0000	undef	undef
	Hoden 0.0000	0.0000	undef	undef
	Lunge 0.0000	0.0000	undef	undef
20	Magen-Speiseroehre 0.0000	0.0000	undef	undef
	Muskel-Skelett 0.0000	0.0000	undef	undef
	Niere 0.0000	0.0000	undef	undef
	Pankreas 0.0380	0.0055	6.8804	0.1453
	Penis 0.0000	0.0000	undef	undef
	Prostata 0.0000	0.0000	undef	undef
25	Uterus_Endometrium 0.0000	0.0000	undef	undef
	Uterus_Myometrium 0.0000	0.0000	undef	undef
	Uterus_allgemein 0.0000	0.0000	undef	undef
	Brust-Hyperplasie 0.0000			
	Prostata-Hyperplasie 0.0000			
	Samenblase 0.0000			
30	Sinnesorgane 0.0000			
	Weisse_Blutkoerperchen 0.0000			
	Zervix 0.0000			
35	FOETUS			
	%Haeufigkeit			
	Entwicklung 0.0000			
	Gastrointenstinal 0.0000			
	Gehirn 0.0000			
40	Haematopoetisch 0.0000			
	Haut 0.0000			
	Hepatisch 0.0000			
	Herz-Blutgefaesse 0.0000			
	Lunge 0.0000			
45	Nebenniere 0.0000			
	Niere 0.0000			
	Placenta 0.0000			
	Prostata 0.0000			
	Sinnesorgane 0.0000			
50	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN			
	%Haeufigkeit			
	Brust 0.0000			
55	Eierstock_n 0.0000			
	Eierstock_t 0.0000			
	Endokrines_Gewebe 0.0000			
	Foetal 0.0000			
60	Gastrointestinal 0.0000			
	Haematopoetisch 0.0000			
	Haut-Muskel 0.0000			
	Hoden 0.0000			
	Lunge 0.0000			
	Nerven 0.0000			
65	Prostata 0.0000			
	Sinnesorgane 0.0000			
	Uterus_n 0.0000			

## Elektronischer Northern für SEQ. ID. NO: 28

	NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse	
			N/T	T/N
5	Blase 0.0000	0.0000	undef	undef
	Brust 0.0000	0.0000	undef	undef
	Duenndarm 0.0000	0.0000	undef	undef
	Eierstock 0.0000	0.0000	undef	undef
10	Endokrines_Gewebe 0.0000	0.0000	undef	undef
	Gastrointestinal 0.0000	0.0000	undef	undef
	Gehirn 0.0000	0.0000	undef	undef
	Haematopoetisch 0.0000	0.0000	undef	undef
	Haut 0.0000	0.0000	undef	undef
15	Hepatisch 0.0000	0.0000	undef	undef
	Herz 0.0000	0.0000	undef	undef
	Hoden 0.0000	0.0000	undef	undef
	Lunge 0.0000	0.0000	undef	undef
20	Magen-Speiseroehre 0.0000	0.0000	undef	undef
	Muskel-Skelett 0.0000	0.0000	undef	undef
	Niere 0.0000	0.0000	undef	undef
	Pankreas 0.0446	0.0000	undef	0.0000
	Penis 0.0000	0.0000	undef	undef
	Prostata 0.0000	0.0000	undef	undef
25	Uterus_Endometrium 0.0000	0.0000	undef	undef
	Uterus_Myometrium 0.0000	0.0000	undef	undef
	Uterus_allgemein 0.0000	0.0000	undef	undef
	Brust-Hyperplasie 0.0000			
	Prostata-Hyperplasie 0.0000			
	Samenblase 0.0000			
30	Sinnesorgane 0.0000			
	Weisse_Blutkoerperchen 0.0000			
	Zervix 0.0000			
35	FOETUS			
	%Haeufigkeit			
	Entwicklung 0.0000			
	Gastrointestinal 0.0056			
	Gehirn 0.0000			
40	Haematopoetisch 0.0000			
	Haut 0.0000			
	Hepatisch 0.0000			
	Herz-Blutgefaesse 0.0000			
	Lunge 0.0000			
45	Nebenniere 0.0000			
	Niere 0.0000			
	Placenta 0.0000			
	Prostata 0.0000			
	Sinnesorgane 0.0000			
50	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN			
	%Haeufigkeit			
	Brust 0.0000			
55	Eierstock_n 0.0000			
	Eierstock_t 0.0000			
	Endokrines_Gewebe 0.0000			
	Foetal 0.0000			
60	Gastrointestinal 0.0000			
	Haematopoetisch 0.0000			
	Haut-Muskel 0.0000			
	Hoden 0.0000			
	Lunge 0.0000			
	Nerven 0.0000			
65	Prostata 0.0000			
	Sinnesorgane 0.0000			
	Uterus_n 0.0000			

## Elektronischer Northern für SEQ. ID. NO: 29

	NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse N/T      T/N
5	Blase 0.0000	0.0000	undef undef
	Brust 0.0000	0.0000	undef undef
	Duenndarm 0.0000	0.0000	undef undef
	Eierstock 0.0000	0.0000	undef undef
10	Endokrines_Gewebe 0.0000	0.0000	undef undef
	Gastrointestinal 0.0000	0.0000	undef undef
	Gehirn 0.0000	0.0000	undef undef
	Haematopoetisch 0.0000	0.0000	undef undef
	Haut 0.0000	0.0000	undef undef
15	Hepatisch 0.0000	0.0000	undef undef
	Herz 0.0000	0.0000	undef undef
	Hoden 0.0000	0.0000	undef undef
	Lunge 0.0000	0.0000	undef undef
20	Magen-Speiseroehre 0.0000	0.0000	undef undef
	Muskel-Skelett 0.0000	0.0000	undef undef
	Niere 0.0000	0.0000	undef undef
	Pankreas 0.0248	0.0000	undef 0.0000
	Penis 0.0000	0.0000	undef undef
	Prostata 0.0000	0.0000	undef undef
25	Uterus_Endometrium 0.0000	0.0000	undef undef
	Uterus_Myometrium 0.0000	0.0000	undef undef
	Uterus_allgemein 0.0000	0.0000	undef undef
	Brust-Hyperplasie 0.0000		
	Prostata-Hyperplasie 0.0000		
	Samenblase 0.0000		
30	Sinnesorgane 0.0000		
	Weisse_Blutkoerperchen 0.0000		
	Zervix 0.0000		

	FOETUS %Haeufigkeit
35	Entwicklung 0.0000
	Gastrointenstinal 0.0000
	Gehirn 0.0000
40	Haematopoetisch 0.0000
	Haut 0.0000
	Hepatisch 0.0000
	Herz-Blutgefaesse 0.0000
	Lunge 0.0000
45	Nebenniere 0.0000
	Niere 0.0000
	Placenta 0.0000
	Prostata 0.0000
50	Sinnesorgane 0.0000

	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN %Haeufigkeit
55	Brust 0.0000
	Eierstock_n 0.0000
	Eierstock_t 0.0000
	Endokrines_Gewebe 0.0000
	Foetal 0.0000
60	Gastrointestinal 0.0122
	Haematopoetisch 0.0000
	Haut-Muskel 0.0000
	Hoden 0.0000
	Lunge 0.0000
	Nerven 0.0000
65	Prostata 0.0000
	Sinnesorgane 0.0000
	Uterus_n 0.0000

## Elektronischer Northern für SEQ. ID. NO: 30

		NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse
				N/T      T/N
5		Blase 0.0000	0.0026	0.0000 undef
		Brust 0.0000	0.0000	undef undef
		Duenndarm 0.0000	0.0000	undef undef
		Eierstock 0.0000	0.0000	undef undef
10	Endokrines_Gewebe	0.0000	0.0000	undef undef
	Gastrointestinal	0.0000	0.0000	undef undef
	Gehirn	0.0000	0.0000	undef undef
	Haematopoetisch	0.0000	0.0000	undef undef
	Haut	0.0000	0.0000	undef undef
15	Hepatisch	0.0000	0.0000	undef undef
	Herz	0.0000	0.0000	undef undef
	Hoden	0.0000	0.0000	undef undef
	Lunge	0.0000	0.0000	undef undef
	Magen-Speiseroehre	0.0000	0.0000	undef undef
	Muskel-Skelett	0.0000	0.0000	undef undef
20	Niere	0.0000	0.0000	undef undef
	Pankreas	0.2693	0.0110	24.3805 0.0410
	Penis	0.0000	0.0000	undef undef
	Prostata	0.0022	0.0000	undef 0.0000
25	Uterus_Endometrium	0.0000	0.0000	undef undef
	Uterus_Myometrium	0.0000	0.0000	undef undef
	Uterus_allgemein	0.0000	0.0000	undef undef
	Brust-Hyperplasie	0.0000		
	Prostata-Hyperplasie	0.0000		
	Samenblase	0.0000		
30	Sinnesorgane	0.0000		
	Weisse_Blutkoerperchen	0.0000		
	Zervix	0.0000		
35		FOETUS %Haeufigkeit		
	Entwicklung	0.0000		
	Gastrointenstinal	0.0028		
	Gehirn	0.0000		
40	Haematopoetisch	0.0000		
	Haut	0.0000		
	Hepatisch	0.0000		
	Herz-Blutgefaesse	0.0000		
	Lunge	0.0000		
45	Nebenniere	0.0000		
	Niere	0.0000		
	Placenta	0.0000		
	Prostata	0.0000		
	Sinnesorgane	0.0000		
50		NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN %Haeufigkeit		
	Brust	0.0000		
55	Eierstock_n	0.0000		
	Eierstock_t	0.0000		
	Endokrines_Gewebe	0.0000		
	Foetal	0.0000		
60	Gastrointestinal	0.0000		
	Haematopoetisch	0.0000		
	Haut-Muskel	0.0000		
	Hoden	0.0000		
	Lunge	0.0000		
	Nerven	0.0000		
65	Prostata	0.0000		
	Sinnesorgane	0.0000		
	Uterus_n	0.0000		

## Elektronischer Northern für SEQ. ID. NO: 31

		NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse N/T	T/N
5	Blase	0.0000	0.0000	undef	undef
	Brust	0.0000	0.0000	undef	undef
	Duenndarm	0.0000	0.0000	undef	undef
	Eierstock	0.0000	0.0000	undef	undef
10	Endokrines_Gewebe	0.0000	0.0000	undef	undef
	Gastrointestinal	0.0000	0.0000	undef	undef
	Gehirn	0.0000	0.0000	undef	undef
	Haematopoetisch	0.0013	0.0000	undef	0.0000
	Haut	0.0000	0.0000	undef	undef
15	Hepatisch	0.0000	0.0000	undef	undef
	Herz	0.0000	0.0000	undef	undef
	Hoden	0.0000	0.0000	undef	undef
	Lunge	0.0000	0.0000	undef	undef
20	Magen-Speiseroehre	0.0000	0.0000	undef	undef
	Muskel-Skelett	0.0000	0.0000	undef	undef
	Niere	0.0000	0.0000	undef	undef
	Pankreas	0.1850	0.0110	16.7523	0.0597
	Penis	0.0000	0.0000	undef	undef
	Prostata	0.0000	0.0000	undef	undef
25	Uterus_Endometrium	0.0000	0.0000	undef	undef
	Uterus_Myometrium	0.0000	0.0000	undef	undef
	Uterus_allgemein	0.0000	0.0000	undef	undef
	Brust-Hyperplasie	0.0000			
	Prostata-Hyperplasie	0.0000			
	Samenblase	0.0000			
30	Sinnesorgane	0.0000			
	Weisse_Blutkoerperchen	0.0000			
	Zervix	0.0000			
35	FOETUS				
	%Haeufigkeit				
	Entwicklung	0.0000			
	Gastrointenstinal	0.0000			
	Gehirn	0.0000			
40	Haematopoetisch	0.0000			
	Haut	0.0000			
	Hepatisch	0.0000			
	Herz-Blutgefaesse	0.0000			
	Lunge	0.0000			
45	Nebenniere	0.0000			
	Niere	0.0000			
	Placenta	0.0061			
	Prostata	0.0000			
	Sinnesorgane	0.0000			
50	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN				
	%Haeufigkeit				
	Brust	0.0000			
55	Eierstock_n	0.0000			
	Eierstock_t	0.0000			
	Endokrines_Gewebe	0.0000			
	Foetal	0.0000			
60	Gastrointestinal	0.0000			
	Haematopoetisch	0.0000			
	Haut-Muskel	0.0000			
	Hoden	0.0000			
	Lunge	0.0000			
	Nerven	0.0000			
65	Prostata	0.0000			
	Sinnesorgane	0.0000			
	Uterus_n	0.0000			

## Elektronischer Northern für SEQ. ID. NO: 35

	NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse N/T      T/N
5	Blase 0.0000	0.0000	undef undef
	Brust 0.0000	0.0000	undef undef
	Duenndarm 0.0000	0.0000	undef undef
	Eierstock 0.0000	0.0000	undef undef
10	Endokrines_Gewebe 0.0000	0.0000	undef undef
	Gastrointestinal 0.0000	0.0000	undef undef
	Gehirn 0.0000	0.0000	undef undef
	Haematopoetisch 0.0000	0.0000	undef undef
	Haut 0.0000	0.0000	undef undef
15	Hepatisch 0.0000	0.0000	undef undef
	Herz 0.0000	0.0000	undef undef
	Hoden 0.0000	0.0000	undef undef
	Lunge 0.0000	0.0000	undef undef
	Magen-Speiseroehre 0.0000	0.0000	undef undef
	Muskel-Skelett 0.0000	0.0000	undef undef
20	Niere 0.0000	0.0000	undef undef
	Pankreas 0.0380	0.0000	undef 0.0000
	Penis 0.0000	0.0000	undef undef
	Prostata 0.0000	0.0000	undef undef
25	Uterus_Endometrium 0.0000	0.0000	undef undef
	Uterus_Myometrium 0.0000	0.0000	undef undef
	Uterus_allgemein 0.0000	0.0000	undef undef
	Brust-Hyperplasie 0.0000		
	Prostata-Hyperplasie 0.0000		
	Samenblase 0.0000		
30	Sinnesorgane 0.0000		
	Weisse_Blutkoerperchen 0.0000		
	Zervix 0.0000		
35	FOETUS		
	%Haeufigkeit		
	Entwicklung 0.0000		
	Gastrointenstinal 0.0083		
	Gehirn 0.0000		
40	Haematopoetisch 0.0000		
	Haut 0.0000		
	Hepatisch 0.0000		
	Herz-Blutgefaesse 0.0000		
	Lunge 0.0000		
45	Nebenniere 0.0000		
	Niere 0.0000		
	Placenta 0.0000		
	Prostata 0.0000		
	Sinnesorgane 0.0000		
50	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN		
	%Haeufigkeit		
	Brust 0.0000		
55	Eierstock_n 0.0000		
	Eierstock_t 0.0000		
	Endokrines_Gewebe 0.0000		
	Foetal 0.0000		
60	Gastrointestinal 0.0000		
	Haematopoetisch 0.0000		
	Haut-Muskel 0.0000		
	Hoden 0.0000		
	Lunge 0.0000		
	Nerven 0.0000		
65	Prostata 0.0000		
	Sinnesorgane 0.0000		
	Uterus_n 0.0000		

## Elektronischer Northern für SEQ. ID. NO: 36

	NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse	
			N/T	T/N
5	Blase 0.0000	0.0026	0.0000	undef
	Brust 0.0000	0.0000	undef	undef
	Duenndarm 0.0000	0.0000	undef	undef
	Eierstock 0.0000	0.0000	undef	undef
10	Endokrines_Gewebe 0.0000	0.0000	undef	undef
	Gastrointestinal 0.0000	0.0000	undef	undef
	Gehirn 0.0000	0.0000	undef	undef
	Haematopoetisch 0.0000	0.0000	undef	undef
	Haut 0.0000	0.0000	undef	undef
15	Hepatisch 0.0000	0.0000	undef	undef
	Herz 0.0000	0.0000	undef	undef
	Hoden 0.0000	0.0000	undef	undef
	Lunge 0.0000	0.0000	undef	undef
	Magen-Speiseroehre 0.0000	0.0000	undef	undef
	Muskel-Skelett 0.0000	0.0000	undef	undef
20	Niere 0.0000	0.0000	undef	undef
	Pankreas 0.3354	0.0166	20.2423	0.0494
	Penis 0.0000	0.0000	undef	undef
	Prostata 0.0022	0.0000	undef	0.0000
25	Uterus_Endometrium 0.0000	0.0000	undef	undef
	Uterus_Myometrium 0.0000	0.0000	undef	undef
	Uterus_allgemein 0.0000	0.0000	undef	undef
	Brust-Hyperplasie 0.0000			
	Prostata-Hyperplasie 0.0000			
	Samenblase 0.0000			
30	Sinnesorgane 0.0000			
	Weisse_Blutkoerperchen 0.0000			
	Zervix 0.0000			

	FOETUS	
	%Haeufigkeit	
35	Entwicklung 0.0000	
	Gastrointenstinal 0.0028	
	Gehirn 0.0000	
40	Haematopoetisch 0.0000	
	Haut 0.0000	
	Hepatisch 0.0000	
	Herz-Biutgefaesse 0.0000	
	Lunge 0.0000	
45	Nebenniere 0.0000	
	Niere 0.0000	
	Placenta 0.0000	
	Prostata 0.0000	
50	Sinnesorgane 0.0000	

	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN	
	%Haeufigkeit	
55	Brust 0.0000	
	Eierstock_n 0.0000	
	Eierstock_t 0.0000	
	Endokrines_Gewebe 0.0000	
	Foetal 0.0000	
60	Gastrointestinal 0.0000	
	Haematopoetisch 0.0000	
	Haut-Muskel 0.0000	
	Hoden 0.0000	
	Lunge 0.0000	
	Nerven 0.0000	
65	Prostata 0.0000	
	Sinnesorgane 0.0000	
	Uterus_n 0.0000	

## Elektronischer Northern für SEQ. ID. NO: 37

	NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse	
			N/T	T/N
5	Blase 0.0000	0.0000	undef	undef
	Brust 0.0000	0.0000	undef	undef
	Duenndarm 0.0000	0.0000	undef	undef
	Eierstock 0.0000	0.0000	undef	undef
10	Endokrines_Gewebe 0.0000	0.0000	undef	undef
	Gastrointestinal 0.0000	0.0000	undef	undef
	Gehirn 0.0000	0.0000	undef	undef
	Haematopoetisch 0.0000	0.0000	undef	undef
	Haut 0.0000	0.0000	undef	undef
	Hepatisch 0.0000	0.0000	undef	undef
15	Herz 0.0000	0.0000	undef	undef
	Hoden 0.0000	0.0000	undef	undef
	Lunge 0.0000	0.0000	undef	undef
	Magen-Speiseroehre 0.0000	0.0000	undef	undef
	Muskel-Skelett 0.0000	0.0000	undef	undef
20	Niere 0.0000	0.0000	undef	undef
	Pankreas 0.0198	0.0000	undef	0.0000
	Penis 0.0000	0.0000	undef	undef
	Prostata 0.0000	0.0000	undef	undef
25	Uterus_Endometrium 0.0000	0.0000	undef	undef
	Uterus_Myometrium 0.0000	0.0000	undef	undef
	Uterus_allgemein 0.0000	0.0000	undef	undef
	Brust-Hyperplasie 0.0000			
	Prostata-Hyperplasie 0.0000			
	Samenblase 0.0000			
30	Sinnesorgane 0.0000			
	Weisse_Blutkoerperchen 0.0000			
	Zervix 0.0000			
35	FOETUS			
	%Haeufigkeit			
	Entwicklung 0.0000			
	Gastrointestinal 0.0000			
	Gehirn 0.0000			
40	Haematopoetisch 0.0000			
	Haut 0.0000			
	Hepatisch 0.0000			
	Herz-Blutgefaesse 0.0000			
	Lunge 0.0000			
45	Nebenniere 0.0000			
	Niere 0.0000			
	Placenta 0.0000			
	Prostata 0.0000			
	Sinnesorgane 0.0000			
50	NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN			
	%Haeufigkeit			
	Brust 0.0000			
55	Eierstock_n 0.0000			
	Eierstock_t 0.0000			
	Endokrines_Gewebe 0.0000			
	Foetal 0.0000			
	Gastrointestinal 0.0000			
60	Haematopoetisch 0.0000			
	Haut-Muskel 0.0000			
	Hoden 0.0000			
	Lunge 0.0000			
	Nerven 0.0000			
65	Prostata 0.0000			
	Sinnesorgane 0.0000			
	Uterus_n 0.0000			

## Elektronischer Northern fuer Seq-ID: 67

		NORMAL %Haeufigkeit	TUMOR %Haeufigkeit	Verhaeltnisse N/T	T/N
5	B_Lymphom	0.0000	0.0000	undef	undef
	Blase	0.0000	0.0000	undef	undef
	Brust	0.0000	0.0000	undef	undef
	Dickdarm	0.0000	0.0000	undef	undef
10	Duenndarm	0.0000	0.0000	undef	undef
	Eierstock	0.0000	0.0000	undef	undef
	Endokrines_Gewebe	0.0000	0.0000	undef	undef
	Gehirn	0.0000	0.0000	undef	undef
	Haut	0.0000	0.0000	undef	undef
15	Hepatisch	0.0000	0.0000	undef	undef
	Herz	0.0000	0.0000	undef	undef
	Hoden	0.0000	0.0000	undef	undef
	Lunge	0.0000	0.0000	undef	undef
20	Magen-Speiseroehre	0.0000	0.0000	undef	undef
	Muskel-Skelett	0.0000	0.0000	undef	undef
	Niere	0.0000	0.0000	undef	undef
	Pankreas	0.1999	0.0055	36.1978	0.0276
	Prostata	0.0000	0.0013	0.0000	undef
	T_Lymphom	0.0051	0.0000	undef	0.0000
	Uterus	0.0000	0.0000	undef	undef
25	Weisse_Blutkoerperchen	0.0014	0.0000	undef	0.0000
	Haematopoetisch	0.0000			
	Penis	0.0000			
	Samenblase	0.0000			
	Sinnesorgane	0.0000			
30			FOETUS %Haeufigkeit		
	Entwicklung	0.0000			
35	Gastrointestinal	0.0000			
	Gehirn	0.0000			
	Haematopoetisch	0.0000			
	Haut	0.0000			
40	Hepatisch	0.0000			
	Herz-Blutgefaesse	0.0000			
	Lunge	0.0000			
	Nebenniere	0.0000			
	Niere	0.0000			
45	Placenta	0.0000			
	Prostata	0.0000			
	Sinnesorgane	0.0000			
			NORMIERTE/SUBTRAHIERTE BIBLIOTHEKEN %Haeufigkeit		
50	Brust	0.0000			
	Brust_t	0.0000			
	Dickdarm_t	0.0000			
	Eierstock_n	0.0000			
	Eierstock_t	0.0000			
55	Endokrines_Gewebe	0.0000			
	Foetal	0.0000			
	Gastrointestinal	0.0244			
	Haematopoetisch	0.0000			
	Haut-Muskel	0.0000			
60	Hoden_n	0.0000			
	Hoden_t	0.0000			
	Lunge_n	0.0000			
	Lunge_t	0.0000			
	Nerven	0.0000			
65	Niere_t	0.0000			
	Ovar_Uterus	0.0000			
	Prostata_n	0.0000			
	Sinnesorgane	0.0000			
	Weisse_Blutkoerperchen	0.0000			

## 2.2 Fisher-Test

Um zu entscheiden, ob eine Partial-Sequenz  $S$  eines Gens in einer Bibliothek für Normal-Gewebe signifikant häufiger oder seltener vorkommt als in einer Bibliothek für entartetes Gewebe, wird Fishers Exakter Test, ein statistisches Standardverfahren (Hays, W. L., (1991) Statistics, Harcourt Brace College Publishers, Fort Worth), durchgeführt.

Die Null-Hypothese lautet: die beiden Bibliotheken können bezüglich der Häufigkeit zu  $S$  homologer Sequenzen nicht unterschieden werden. Falls die Null-Hypothese mit hinreichend hoher Sicherheit abgelehnt werden kann, wird das zu  $S$  gehörende Gen als interessanter Kandidat für ein Krebs-Gen akzeptiert, und es wird im nächsten Schritt versucht, eine Verlängerung seiner Sequenz zu erreichen.

## 15 Beispiel 3

### 20 Automatische Verlängerung der Partial-Sequenz

Die automatische Verlängerung der Partial-Sequenz  $S$  vollzieht sich in drei Schritten:

1. Ermittlung aller zu  $S$  homologen Sequenzen aus der Gesamtmenge der zur Verfügung stehenden Sequenzen mit Hilfe von BLAST
2. Assemblierung dieser Sequenzen mittels des Standardprogramms GAP4 (Bonfield, J. K., Smith, K. F., und Staden R. (1995), Nucleic Acids Research 23 4992-4999) (Contig-Bildung).
- 30 3. Berechnung einer Konsens-Sequenz  $C$  aus den assemblierten Sequenzen

35 Die Konsens-Sequenz  $C$  wird im allgemeinen länger sein als die Ausgangssequenz  $S$ . Ihr elektronischer Northern-Blot wird demzufolge von dem für  $S$  abweichen. Ein erneuter Fisher-Test entscheidet, ob die Alternativ-Hypothese der Abweichung von einer gleichmäßigen Expression in beiden Bibliotheken aufrechterhalten werden kann. Ist dies der Fall, wird versucht,  $C$  in gleicher Weise wie  $S$  zu verlängern. Diese Iteration wird mit der jeweils erhaltenen Konsensus-Sequenzen  $C_i$  ( $i$ : Index der 40 Iteration) fortgesetzt, bis die Alternativ-Hypothese verworfen wird (if  $H_0$  Exit; Abbruchkriterium I) oder bis keine automatische Verlängerung mehr möglich ist (while  $C_i > C_{i-1}$ ; Abbruchkriterium II).

45 Im Fall des Abbruchkriteriums II bekommt man mit der nach der letzten Iteration vorliegenden Konsens-Sequenz eine komplett oder annähernd komplett Sequenz eines Gens, das mit hoher statistischer Sicherheit mit Krebs in Zusammenhang gebracht werden kann.

50 Analog der oben beschriebenen Beispiele konnten die in der Tabelle I beschriebenen Nukleinsäure-Sequenzen aus Pankreasnormalgewebe gefunden werden.

Ferner konnten zu den einzelnen Nukleinsäure-Sequenzen die Peptidsequenzen (ORF's) bestimmt werden, die in der Tabelle II aufgelistet sind, wobei wenigen Nukleinsäure-Sequenzen kein Peptid zugeordnet werden kann und einigen Nukleinsäure-Sequenzen mehr als ein Peptid zugeordnet werden kann. Wie bereits oben erwähnt, sind sowohl die ermittelten Nukleinsäure-Sequenzen, als auch die den Nukleinsäure-Sequenzen zugeordneten Peptid-Sequenzen Gegenstand der vorliegenden Erfindung.

#### Beispiel 4

##### **Kartierung der Nukleinsäure-Sequenzen auf dem humanen Genom**

Die Kartierung der humanen Gene erfolgte unter Verwendung des Stanford G3 Hybrid-Panels (Stewart et al., 1997), der von Research Genetics, Huntsville, Alabama vertrieben wird. Dieses Panel besteht aus 83 verschiedenen genomischen DNAs von Mensch-Hamster Hybridzelllinien und erlaubt eine Auflösung von 500 Kilobasen. Die Hybridzelllinien wurden durch Fusion von bestrahlten diploiden menschlichen Zellen mit Zellen des Chinesischen Hamsters gewonnen. Das Rückhaltemuster der humanen Chromosomenfragmente wird mittels genspezifischer Primer in einer Polymerase-Kettenreaktion bestimmt und mit Hilfe der vom Stanford RH Server verfügbaren Software analysiert ([http://www.stanford.edu/RH/rhserver\\_form2.html](http://www.stanford.edu/RH/rhserver_form2.html)). Dieses Programm bestimmt den STS-Marker, der am nächsten zum gesuchten Gen liegt. Die entsprechende zytogenetische Bande wurde unter Verwendung des "Mapview" -Programms der Genome Database (GDB), (<http://gdbwww.dkfz-heidelberg.de>) bestimmt. Neben dem kartieren von Genen auf dem menschlichen Chromosomensatz durch verschiedene experimentelle Methoden ist es möglich die Lage von Genen auf diesem durch bioinformatische Methoden zu bestimmen. Dazu wurde das bekannte Programm e-PCR eingesetzt (Schuler GD (1998) Electronic PCR: bridging the gap between genome mapping and genome sequencing. Trends Biotechnol 16; 456-459, Schuler GD (1997). Sequence mapping by electronic PCR. Genome Res 7; 541-550). Die dabei eingesetzte Datenbank entspricht nicht mehr der in der Literatur angegebenen, sonder ist eine Weiterentwicklung, welche Daten der öffentlichen Datenbank RHdb (<http://www.ebi.ac.uk/RHdb/index.html>) einschließt. Analog zu der Kartierung durch die Hybrid-Panels erfolgte eine Auswertung der Ergebnisse mit der obengenannten Software und der Software des Whitehead-Institutes (<http://carbon.wi.mit.edu:8000/cgi-bin/contig/rhmapper.pl>).

**Beispiel 5****Gewinnung von genomischen DNA-Sequenzen (BAC-Klone)**

5 Die die entsprechenden cDNA enthaltenen genomischen BAC-Klone (<http://www.tree.caltech.edu/>; Shizuya, H., B. Birren, U-J. Kim, V. Mancino, T. Slepak, Y. Tachiiri, M. Simon (1992) Proc. Natl. Acad. Sci., USA 89: 8794-8797) wurden mit der Prozedur des "down-to-the-well" isoliert. Bei dieser Prozedur wird eine Bibliothek bestehend aus BAC-Klonen (die Bibliothek überdeckt ca. 3 x das humane Genom)

10 in ein bestimmtes Raster gebracht, so daß die DNA dieser Klone mit einer spezifischen PCR untersucht werden kann. Dabei erfolgt ein "Poolen" der DNA verschiedener BAC-Klone. Durch eine kombinatorische Analyse ist es möglich die Klone zu bestimmen, die die gesuchte DNA enthalten. Durch das Festlegen der Klone kann die Adresse der Klone in der Bibliothek bestimmt werden. Diese Adresse

15 zusammen mit dem Namen der verwendeten Bibliothek legen die Klone und damit die DNA-Sequenz dieser Klone eindeutig fest.

TABELLE I

Seq ID No.	Expression	Funktion	Module	Cytogenetische Lokalisation	Nächster Marker
2	In Pankreasnormalgewebe überexprimiert	Human mRNA for regenerating protein 1 beta	C_TYPE_LEC TIN_2	2p13.1	CHLC.GCT1B4
4	In Pankreasnormalgewebe überexprimiert	Glanduläres Kallikrein 1	trypsin		
6	In Pankreasnormalgewebe überexprimiert	cDNA encoding human phospholipase A2	phoslip	12q24.22-q24.23 3q26.33-q28	SHGC-10488- AFMa225xe5 D3S1262-D3S1580
7	In Pankreasnormalgewebe überexprimiert	Human somatostatin 1			
9	In Pankreasnormalgewebe überexprimiert	Human islet amyloid protein	Calc_CGRP_1 APP	12p12.3-p12.2	D12S364-D12S310
11	In Pankreasnormalgewebe überexprimiert	Human pancreatic secretory trypsin inhibitor (PSTI)	kazal	12p12.3	WI-7377-D12S1669
12	In Pankreasnormalgewebe überexprimiert	Human protease E	trypsin		
14	In Pankreasnormalgewebe überexprimiert	Homolog zu Trypsin precursor	TRYPSIN_CA TAL Colipase	16q24.1-q24.2 6p21.2-p21.31	CHLC:GATA71F09 D6S439-D6S291
15	In Pankreasnormalgewebe überexprimiert	Humanes Chymotrypsinogen			
16	In Pankreasnormalgewebe überexprimiert	Homo sapiens colipase (CLPS)			
17	In Pankreasnormalgewebe überexprimiert	caldecrin=serum calcium-decreasing factor	TRYPSIN_CA TAL	1p36.13	AFMa127zc9- AFMa232zb9 WI-3768
18	In Pankreasnormalgewebe überexprimiert	Human pancreatic zymogen granule membrane protein GP-2	zona_pellucida	16p11.2	
19	In Pankreasnormalgewebe überexprimiert	Humanes IAPP	7im_2	12p12.3	WI-7377-D12S1669
21	In Pankreasnormalgewebe überexprimiert	H.sapiens mRNA for chymotrypsin-like protease CTRL-1	PRO_RICH	16q22.3-q23.1	WI-3392
24	In Pankreasnormalgewebe überexprimiert	Humanes Homolog zu D. melanogaster Calbindin-32	EF_HAND_2		
25	In Pankreasnormalgewebe überexprimiert	Humanes Homolog zu R. norvegicus syncolin		19q13.2	SHGC-8810-SHGC- 37129
27	In Pankreasnormalgewebe überexprimiert	Mus musculus unknown protein precursor	CUB; zona_pellucida		

Seq ID No.	Expression	Funktion	Module	Cyogenetische Lokalisation	Nächster Marker
28	in Pankreasnormalgewebe überexprimiert	unbekannt		16p12.3-p12.2	D16S3045-SHGC-6042
29	in Pankreasnormalgewebe überexprimiert	unbekannt		16p12.3-p12.2	D16S3045-SHGC-6042
30	in Pankreasnormalgewebe überexprimiert	unbekannt		16p12.3-p12.2	D16S3045-SHGC-6042
31	in Pankreasnormalgewebe überexprimiert	unbekannt	trypsin	16q23.1	SHGC-14629-SHGC-8141
35	in Pankreasnormalgewebe überexprimiert	unbekannt	PRO_RICH	7q32.3-q33	SHGC-30423-AFM183xe11
36	in Pankreasnormalgewebe überexprimiert	unbekannt	PRO_RICH	16p12.3-p12.2	D16S3045-SHGC-6042
37	in Pankreasnormalgewebe überexprimiert	unbekannt		10q25.1-q25.3	AFM249v1-SHGC-14812
67	in Pankreasnormalgewebe überexprimiert	Verlängerung zu Seq ID: 27			

## TABELLE II

DNA Seq ID No:	Peptid Seq ID No:
14	68
	69
24	39
	40
25	41
	42
	43
27	44
28	45
29	46
30	47
	48
	49
31	50
	51
	52
35	56
	57
	58
36	59
	60
	61
37	62
	63
67	70
	71

5 Die erfinderischen Nukleinsäure-Sequenzen Seq. ID No. 1 bis Seq. ID No. 37 und 67 der ermittelten Kandidatengene und die ermittelten Aminosäure-Sequenzen Seq. ID No. 39 bis Seq. ID No. 63 und 68-71 werden in dem nachfolgenden Sequenzprotokoll beschrieben.

**Sequenzprotokoll****(1) ALLGEMEINE INFORMATION:**

5 (i) ANMELDER:  
(A) NAME: metaGen - Gesellschaft für Genomforschung mbH  
(B) STRASSE: Ihnestrassse 63  
(C) STADT: Berlin  
(E) LAND: Deutschland  
10 (F) POST CODE (ZIP): D-14195  
(G) TELEFON: (030)-8413 1673  
(H) TELEFAX: (030)-8413 1674

15 (ii) TITEL DER ERFINDUNG: Menschliche Nukleinsäure-Sequenzen aus  
Pankreasnormalgewebe

20 (iii) Anzahl der Sequenzen: 51  
(iv) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

**(2) INFORMATION ÜBER SEQ ID NO:2:**

30 (i) SEQUENZ CHARAKTERISTIK:  
(A) LÄNGE: 836 Basenpaare  
(B) TYP: Nukleinsäure  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear  
35 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
hergestellte partielle cDNA  
(iii) HYPOTHETISCH: NEIN  
40 (iii) ANTI-SENSE: NEIN  
(vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH  
(C) ORGAN:  
45 (vii) SONSTIGE HERKUNFT:  
(A) BIBLIOTHEK: cDNA library  
(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:2

50 ctcgtgcgaa ttcggcagaa ctctgtctga cctgacaaggc cacctcaagt ggacaaggca 60  
cttaccaaca gagattgctg atttgctcct taagaagag attcactgccc gctaaggatg120  
gctcagacca actcggtctt catgctgatc tcctccctga tgttcctgtc tctgagccaa180

5 ggccaggagt cccagacaga gctgcctaat ccccgaatca gctgcccaga aggcaccaat240  
 gcctatcgct cctactgcta ctactttaat gaagaccctg agacactgggt tcatgcagat300  
 ctctattgcc agaacatgaa ttcaaggcaac ctgggtctg tgcgcaccca ggcggagggt360  
 gccttcgtgg cctcactgat taaggagagt agcactgatg acagcaatgt ctggattggc420  
 10 ctccatgacc caaaaaagaa ccggcgtgg cactggagta gtgggtccct ggtctcctac480  
 aagtcttggg acactggat cccgagcagt gctaattgtg gctactgtgc aagccgtact540  
 tcatgctcag gattcaagaa atgaaaggat gaatcttgtg agaagaaggt ctcccttgtt600  
 tgcaagttca aaaactagag gaagctgaaa aatggatgtc tagaactgggt cctgcaatta660  
 ctatgaagtc aaaaattaaa cttagactatg tctccaactc agttcagacc atctcctccc720  
 taatgagttt gcatcgctga tcttcagttc cttcacctgt ctcagtctct agagccctga780  
 aaaataaaaa caaacttatt tttatccagt gaacaaagag agggaggaga agagac 836

(2) INFORMATION ÜBER SEQ ID NO: 4:

15 (i) SEQUENZ CHARAKTERISTIK:

- (A) LÄNGE: 871 Basenpaare
- (B) TYP: Nukleinsäure
- (C) STRANG: einzel
- (D) TOPOLOGIE: linear

20 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
 hergestellte partielle cDNA

(iii) HYPOTHETISCH: NEIN

25 (iii) ANTI-SENSE: NEIN

(vi) HERKUNFT:

(A) ORGANISMUS: MENSCH  
 30 (C) ORGAN:

(vii) SONSTIGE HERKUNFT:

(A) BIBLIOTHEK: cDNA library

35 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:4

40 acctgctggc ccctggacac ctctgtcacc atgtggttcc tggttctgtg cctcgccctg 60  
 tccttgggg ggactggc tgcgcggcc attcagtccc ggattgtgg aggctgggag120  
 tggagcagc attcccaagcc ctggcaggcg gctctgtacc atttcagcac tttccagtgt180  
 gggggcatcc tggtgacccg ccagtgggtg ctcacagctg ctcattgtcat cagcgacaat240  
 taccagctt ggctgggtcg ccacaacttg ttgacgacg aaaacacacgc ccagttgtt300  
 45 catgtcagtg agagcttccc acaccctggc ttcaacatga gcctccttggaa gaaccacacc360  
 cggcaagcag acgaggacta cagccacgac ctcatgtgc tccgcctgac agagcctgct420  
 gataccatca cagacgctgt gaaggtcgtg gagttgccc cccaggaacc cgaagtgggg480  
 50 agcacctgtt tggcttccgg ctggggcagc atcgaaccag agaatttctc attccagat540  
 gatctccagt gtgtggaccc caaaatctt cctaattgtg agtgcgaaaa agcccacgtc600  
 cagaagggtga cagacttcat gctgtgtgc ggacacctgg aaggtggcaa agacacctgt660  
 gtgggtgatt cagggggccc gctgtgtgt gatgggtgtc tccaaagggtgt cacatcatgg720  
 ggctacgtcc cttgtggcac ccccaataag ctttctgtcg ccgtcagagt gctgtcttat780  
 gtgaagtggc tcgaggacac catagcggag aactcctgaa cggccagccc tgccttac840  
 55 ccccaatgtg catccaaaaa a 871

(2) INFORMATION ÜBER SEQ ID NO: 6:

55 (i) SEQUENZ CHARAKTERISTIK:

- (A) LÄNGE: 644 Basenpaare
- (B) TYP: Nukleinsäure

(C) STRANG: einzel  
(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
5 hergestellte partielle cDNA

(iii) HYPOTHETISCH: NEIN

(iii) ANTI-SENSE: NEIN

10 (vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH  
(C) ORGAN:

15 (vii) SONSTIGE HERKUNFT:  
(A) BIBLIOTHEK: cDNA library

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:6

20 ctgactataa gactataacct gagactggtc atctcagttc ttttctcacc ttgactgcaa 60  
gatgaaactc cttgtgtctag ctgtgtgtc cacagtggcc gcccggaca gcggcatcag120  
ccctcgggcc gtgtggcagt tccgaaaaat gatcaagtgc gtgatcccgg ggagtgaccc180  
cttcttggaa tacaacaact acggctcta ctgtggcttg gggggctcag gcacccccgt240  
ggatgaactg gacaagtgtc gccagacaca tgacaactgc tatgaccagg ccaagaagct300  
25 ggacagctgt aaattttgtc tgacacaaccc gtacacccac acctattcat actcgtgctc360  
tggctcggca atcacctgtc gcaagaaaaaa caaagagtgt gaggccttca tttgcaactg420  
cgaccgcaac gctgccatct gctttcaaa agctccatat aacaaggcac acaagaacct480  
ggacaccaag aagtattgtc agagttgaat atcacctctc aaaagcatca cctctatctg540  
30 cctcatctca cactgtactc tccaataaaag cacctgttg aaagacaaaa gaaaaaaaaa600  
gaaaaaaaaaaga aaaaagaaaag gaaagggaaag cagggaaaat ggaa 644

(2) INFORMATION ÜBER SEQ ID NO: 7:

(i) SEQUENZ CHARAKTERISTIK:  
35 (A) LÄNGE: 723 Basenpaare  
(B) TYP: Nukleinsäure  
(C) STRANG: einzel  
(D) TOPOLOGIE: linear

40 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
hergestellte partielle cDNA

(iii) HYPOTHETISCH: NEIN

45 (iii) ANTI-SENSE: NEIN

(vi) HERKUNFT:  
50 (A) ORGANISMUS: MENSCH  
(C) ORGAN:

(vii) SONSTIGE HERKUNFT:  
(A) BIBLIOTHEK: cDNA library

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:7

attttgcgag	gctaatggtg	cgtaaaaagg	ctggtgagat	ctgggggcgc	ctccctagcct	60
gacgtcagag	agagggaggg	gtttaaaaaca	gagggagagc	gttgagagca	cacaaggcgc	120
tttaggagcg	aggttcggag	ccatcgctgc	tgcctgctga	tccgcgccta	gagttgacc	180
agccactctc	cagctcggt	ttcgcggcgc	cgagatgtg	tcctgcccgc	tccagtgcgc	240
gctggctgcg	ctgtccatcg	tcctggccct	gggctgtgtc	accggcgctc	cctcggaccc	300
cagactccgt	cagtttctgc	agaagtccct	ggctgtgc	gcggggaaagc	aggaactggc	360
caagtacttc	ttggcagagc	tgctgtctga	acccaaccag	acggagaatg	atgcccttga	420
acctgaagat	ctgtcccagg	ctgctgagca	ggatgaaatg	aggcttgagc	tgcagagatc	480
tgcttaactca	aaccggcta	tggcaccccg	agaacgc当地	gctggctgca	agaatttctt	540
ctggaaagact	ttcacatcct	gttagcttc	ttaacttagta	ttgtccat	cagacctctg	600
atccctcgcc	cccacacccc	atctctttc	cctaattcctc	caagtcttca	gcgagaccct	660
tgcattagaa	actgaaaact	gtaaatacaa	aataaaaatta	tggtaaaatt	atgagaaaaga	720
aaa						723

723

(2) INFORMATION ÜBER SEQ ID NO: 9:

### (i) SEQUENZ CHARAKTERISTIK:

(A) LÄNGE: 801 Basenpaare

**(B) TYP: Nukleinsäure**

- (C) STRANG: einzel
- (D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung hergestellte partielle cDNA

## 25 hergestellte partielle cDNA

(iii) HYPOTHETISCH: NEIN

(iii) ANTI-SENSE: NEIN

30 (vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH  
(C) ORGAN:

35 (vii) SONSTIGE HERKUNFT:  
(A) BIBLIOTHEK: cDNA library

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:9

```

40  gtcagagctg agaaagggtgt gaggggtata taagagctgg attactagtt agcaaatgag 60
    ggggtaaata ttccagtggc tacaagcttg gactctttc ttgaagcttt ctttctatca120
    gaagcatttg ctgatattgc tgacattgaa acattaaaag aaaatttgag aagcaatggg180
    catcctgaag ctgcaagtat ttctcattgt gctctctgtt gcattgaacc atctgaagc240
    tacaccatt gaaagtcatc aggtggaaaa gcggaaatgc aacactgcca catgtgcaac300
45  gcagcgctg gcaaattttt tagttcattc cagcaacaac ttttgtgcca ttctctcatc360
    taccaacgtg gnatccaata catatggcaa gaggaatgca gtagaggttt taaagagaga420
    gccactgaat tactgcccc ttagaggac aatgtaaactc tatagtttt gtttatgtt480
    ctagtgattt cctgtataat ttaacagtgc cctttcattc tccagtgtga atatatggtc540
    tgtgtgtctg atgtttgtt ctaggacata taccttctca aaagattgtt ttatatgtag600
    tactaactaa ggtcccataaa taaaaagata gtatcttta aatgaaaatg ttttgctat660
    agatttgtat tttaaaacat aagaacgtca ttttgggacc tatatctcag tggcacaggt720
    ttaagaacga aggagaaaaaa ggtagttgga cccttggaaag gtgtgcgcgg ctgatgtgtt780
    gtccgcgggg cggaaatccgg c 801

```

801

55 (2) INFORMATION ÜBER SEQ ID NO: 11:

### (i) SEQUENZ CHARAKTERISTIK:

- (A) LÄNGE: 608 Basenpaare
- (B) TYP: Nukleinsäure
- (C) STRANG: einzeln
- (D) TOPOLOGIE: linear

5

(ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung hergestellte partielle cDNA

10 (iii) HYPOTHETISCH: NEIN

10 (iii) ANTI-SENSE: NEIN

15 (vi) HERKUNFT:

- (A) ORGANISMUS: MENSCH
- (C) ORGAN:

15 (vii) SONSTIGE HERKUNFT:

- (A) BIBLIOTHEK: cDNA library

20

20 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:11

gaaagaaaact ggttgttca ttcctggctg gagagagtag aaaagagaat gaagatggag 60  
taacatacat ttgtgacatt cccagaacct ggaggccagg ctatgacaca gagtcaatca120  
ataaccaggg agatctgtga tatagccag tagtggggc cttgctgcca tctgccccat180  
25 gacccttcca gtcccaggtct tctgaagaga cgtggtaagt ggggtgcagt tttcaactga240  
cctctggacg cagaacttca gccatgaagg taacaggcat ctttcttctc agtgccttgg300  
ccctgtttag tctatctgtt aacactggag ctgactccct gggaaagagag gccaaatgtt360  
acaatgaact taatggatgc accaagatat atgaccctgt ctgtgggact gatgaaata420  
cttatcccaa tgaatgcgtg ttatgttttggaaatcgaa acgcccagact tctatcctca480  
30 ttcaaaaatc tgggccttgc tgagaaccaa gttttgaaa tcccatcagg tcaccgcgag540  
gcctgactgg ccttattgtt gaataaatgt atctgaatat caaaaaaaaaa aaagcgaaaa600  
608  
aagaggaa

35 (2) INFORMATION ÜBER SEQ ID NO: 12:

35

(i) SEQUENZ CHARAKTERISTIK:

- (A) LÄNGE: 892 Basenpaare
- (B) TYP: Nukleinsäure
- (C) STRANG: einzeln
- (D) TOPOLOGIE: linear

40 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung hergestellte partielle cDNA

45

45 (iii) HYPOTHETISCH: NEIN

(iii) ANTI-SENSE: NEIN

50 (vi) HERKUNFT:

- (A) ORGANISMUS: MENSCH
- (C) ORGAN:

(vii) SONSTIGE HERKUNFT:

- (A) BIBLIOTHEK: cDNA library

## (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:12

5           gtcttctcct cttttttgc ttttttctt ggaattttct gagagatctt tattctttat 60  
 tcaggatgtg ggatcgatca gcactgccag ctgggccttg gttctagtg 120  
 tctcctcaat ccagtcgtg aaggcggaga ctcgagtggaa caccgtggc ttcctgcggg180  
 tggtgcagcc aaaggcagaa acaaagctgg tcacgcctat gacgtgcctg 240  
 ctgtgggca gttgagggt cctccagagt caccgttgc gccggagcgg atgtcccctc300  
 cagcacacac catggcttc ttcacggagg aaaaaaccca gttccacctg gagcagtgtt360  
 10          catagtcac cacgggcagc agggctct gcagcttgc tgggagtggc ccgttgat 420  
 agagacggcc ccagccgtg atgtacggg gtgtctcg 480  
 gagggagtga ggcgagctgg acggcgttc ccagctggc gctgcgtga 540  
 gggcgtatgtc attgcccacag gccacacacg agcgggttca gatggatgc acaaagagg 600  
 15          ccccaggtt gatggggatc acctgctcg ggcctctt cacagcacgg tcgtactcgc660  
 ccaacaccac ctggtaggtc caggagctcg agatgcgtg gccggcgtc acaaccagg 720  
 cggggcgtat gaggctaccg ccacacgtgt ggtagaagtt tccacttttc tcataactgca780  
 gggaaacctg ccagggcag ctgttagggg 840  
 aaggcggaga ggaagggtggg ccatacgctg aggcaacggc cacaaggagc gg       892

## 20 (2) INFORMATION ÜBER SEQ ID NO: 14:

## (i) SEQUENZ CHARAKTERISTIK:

- (A) LÄNGE: 229 Basenpaare
- (B) TYP: Nukleinsäure
- 25           (C) STRANG: einzel
- (D) TOPOLOGIE: linear

30           (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung hergestellte partielle cDNA

35           (iii) HYPOTHETISCH: NEIN

40           (iv) ANTI-SENSE: NEIN

## (vi) HERKUNFT:

- (A) ORGANISMUS: MENSCH
- (C) ORGAN:

## (vii) SONSTIGE HERKUNFT:

- 40           (A) BIBLIOTHEK: cDNA library

## (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:14

45           gcagtaatca acgcccgcgg tccaccactc ctgcccaccc ccccccagcc actggcacg 60  
 agcatctcg gctggcaac actgcgaccc gcccggacta cccagacagc gcaggcctg 120  
 tgctccggcg accagcaagt gtgaacctcc taccggaaa gattaccaca acatgttccg 180  
 tggggctcc ttagggcaa ggacacatgt cagtgattct ggtgcctg       229

## (2) INFORMATION ÜBER SEQ ID NO: 15:

## (i) SEQUENZ CHARAKTERISTIK:

- (A) LÄNGE: 885 Basenpaare
- (B) TYP: Nukleinsäure
- 55           (C) STRANG: einzel
- (D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
hergestellte partielle cDNA

(iii) HYPOTHETISCH: NEIN

5

(iii) ANTI-SENSE: NEIN

(vi) HERKUNFT:

10

(A) ORGANISMUS: MENSCH

(C) ORGAN:

(vii) SONSTIGE HERKUNFT:

(A) BIBLIOTHEK: cDNA library

15

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:15

cgcaagcgca tggcttcctt ctggctcctc tcctgcttct cccttgtggg ggccgccttt 60  
ggctgcgggg tccccgcatt ccacccctgtg ctcagcggcc tgcaggat cgtaatggg120  
gaggacgccc tccccggctc ctggccctgg caggtgtccc tgcaggacaa aacccgcttc180  
20 cacttctgctg ggggctccctt catcagcgg gactgggtgg tcaccgctgc ccactgcggg240  
gtcaggaccc tccgacgtggt cgtggctggg gagttgacc agggctctga cgaggagaac300  
atccagggtcc tgaagatcgc caaggtcttc aagaacccca agttcagcat tctgaccgtg360  
aacaatgaca tcaccctgtctt gaagctggcc acacctgccc gcttctccca gacagtgtcc420  
25 gccgtgtgcc tgcccagcgc cgacgacgac ttccccgggg ggacactgtg tgccaccaca480  
ggctggggca agaccaagta caacgcacaa aagacccctg acaagctgca gcaggcagcc540  
ctgcccctcc tgcataatgc cgaatgcaag aagtcttggg gcaggagat caccgacgtg600  
atgatctgtg cggggccag tggcgtctcc tcctgcatgg gcaactctgg cggcccccgtg660  
gtctccaaa agatggagc ctggaccctgg gtgggcattt tgccttgggg cagcgcacacc720  
30 tgctccaccc tccagccctgg cgtgtacgc cgtgtaccca agtcataacc ttgggtgcag780  
aagatcctgg ctgccaactg agcccgccgc tccctccgac cctgctcccc acagagcctc840  
agtaaaccctt tggaaacacaa aaaaaaggaa ggaaaaaaaaa acata 885

(2) INFORMATION ÜBER SEQ ID NO: 16:

35

(i) SEQUENZ CHARAKTERISTIK:

- (A) LÄNGE: 656 Basenpaare
- (B) TYP: Nukleinsäure
- (C) STRANG: einzeln
- (D) TOPOLOGIE: linear

40

(ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
hergestellte partielle cDNA

(iii) HYPOTHETISCH: NEIN

45

(iii) ANTI-SENSE: NEIN

(vi) HERKUNFT:

50

(A) ORGANISMUS: MENSCH

(C) ORGAN:

(vii) SONSTIGE HERKUNFT:

(A) BIBLIOTHEK: cDNA library

55

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:16

5 ggatctcgta aggtcagggc ccctttata gccccatcac cacagctggc tctgtctccc 60  
gccacccaca ccagctgtcc cactcaccat ggagaagatc ctgatcctcc tgcttgcgc120  
cctctctgtg gcctatgcag ctccatggccc cgggggatc attatcaacc tggagaacgg180  
tgagctctgc atgaatagtg cccagtgtaa gagcaattgc tgccagcatt caagtgcgt240  
gggcctggcc cgctgcacat ccatggccag cgagaacagc gagtgcctg tcaagacgct300  
ctatgggatt tactacaagt gtcctgtga gctggcctg acctgtgagg gagacaagac360  
catcggtggc tccatcacca acaccaactt tggcatctgc catgacgctg gacgctccaa420  
10 gcagtgagac tgcccacca ctcccacacc tagccccagaa tgctgttaggc cactaggcgc480  
aggggcatct ctcccctgt ccagcgcatt tcccggtctg gccacctcc tgaccagcat540  
atctgtttc tgattgcgt ctccacaatt aaaggcctcc tgcaaaccctt taaaaaaaaa600  
aagagaagaa gaaaaaaaaa agaaaaaaga agaaaggagg agaagagaag gaggaa 656

(2) INFORMATION ÜBER SEQ ID NO: 17:

15

(i) SEQUENZ CHARAKTERISTIK:  
(A) LÄNGE: 105 Basenpaare  
(B) TYP: Nukleinsäure  
(C) STRANG: einzel  
20 (D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
hergestellte partielle cDNA

25

(iii) HYPOTHETISCH: NEIN

(iv) ANTI-SENSE: NEIN

(v) HERKUNFT:  
30 (A) ORGANISMUS: MENSCH  
(C) ORGAN:

(vi) SONSTIGE HERKUNFT:  
(A) BIBLIOTHEK: cDNA library

35

(vii) SEQUENZ-BESCHREIBUNG: SEQ ID NO:17

tgtctgcgt gctccagcgc atctgccggg ctggccaccc ccttggcggc catatctggg 60  
ttctgattgc gctttcaca attaaaggcg tcctgcagac cttaa 105

40

(2) INFORMATION ÜBER SEQ ID NO: 18:

45

(i) SEQUENZ CHARAKTERISTIK:  
(A) LÄNGE: 1746 Basenpaare  
(B) TYP: Nukleinsäure  
(C) STRANG: einzel  
(D) TOPOLOGIE: linear

50

(ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
hergestellte partielle cDNA

(iii) HYPOTHETISCH: NEIN

(iv) ANTI-SENSE: NEIN

55

(vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH  
(C) ORGAN:

5 (vii) SONSTIGE HERKUNFT:  
(A) BIBLIOTHEK: cDNA library

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:18

10	ctgtgaccag ccccgcccgag gcttagtgcct gcaggaacat tctggaggtt aatggcttc acaccactgc tgccagaagg tatgtttgg gctgatggg cctatcatag tagctcttag 120 atcttggaca agccacttag cttccctgag ccccccagggt tgggtgagag tgcacaagg 180 ggtttacata aagagcttag tataagtctg acatgtgagt ttccataaaacg taagcgacaa 240 ttacgataat ttcaatttagg gatgggggg gaggatgca cagagaagca agcttgggg 300 ttcctgactc ttcccaaggg tcaataattt atcggtctag tttttgcct gtttcctca 360 ttcgccttcc cataaaagtgt cagtaacact gaaggaagag gtcgcctgac ctccatgagc 420 tcagaaggta ggggtggctgc atggtaagc aaattgactc tggaaatcag atctggatt 480 gagatacacc tacatacaca tacacataaa catctgtatg gttttttt tttgccaaag 540 agaaatcaaa cccatgccat ctacaaaaac accctctcct tggtcaatga tttcatcatc 600 agagacacca tcctcaacat caacttccaa tggccttacc cactgacat gaaagtca 660 ctccaagctg ccttgcagcc cattgttaagt tccctgaacg tcagttggg cgggaatgg 720 gagttcattt tcaggatggc cctcttccaa gaccagaact acacgaatcc ttacgaagg 780 gatgcagttt aactgtctgt tggatccgtg ctgtatgtgg gtgcctat tggaaatgg 840 gacacctccc ggttaacact ggtgttggg aactgtatg ccaccccccac tggaaatgg 900 25 gctgacccctt tgaagtattt catcatcaga aacagctgct cttttttt tggatccacc 960 atccacgtgg aggagaatgg gcagtcctcg gaaagccggg tctcaatttca gatgttcatg 1020 tttgctggac attatgaccc ttttttttccctt cattgtgaga ttcatctctg tgattctt 1080 aatgaacactt gccagccccc ttgtcaaga agtcaagtcc gcagttggg tggaaatgg 1140 gacccatggcc ggggtctaga ttttttttccctt atcaactcgaa gaggtgcaca gtctcccggt 1200 30 gtcatgaatg gaaacccctag cactgcaggg ttttttttccctt ctttttttccctt ggtccctcg 1260 actgtccccc tggctggct gtttttttccctt ctttttttccctt ctttttttccctt tggatttgtt 1320 gttcccccctt ctggcaatgg ctttttttccctt ctttttttccctt ctttttttccctt tggatttgtt 1380 gcttggattt aacagaatca aggccaggct aggttaggaa aaggaaatgg ctttcacccctt 1440 ctttaaaactt ctggctggg cgcagttggct catgcctgtt atccacacat tttggggaggc 1500 35 tgaggcagggt ggatcacctg aggtcagcag ttttttttccctt ctttttttccctt ggtccctcg 1560 ctccgtctt actaaaaata ctttttttccctt ctttttttccctt ctttttttccctt tggatttgtt 1620 agctactcg gaggccaaagg ctttttttccctt ctttttttccctt ctttttttccctt tggatttgtt 1680 agttgagattt gtgccattgc actccacccctt ctttttttccctt ctttttttccctt ggtccctcg 1740 aaaaaaa
----	--

40 (2) INFORMATION ÜBER SEQ ID NO: 19:

(i) SEQUENZ CHARAKTERISTIK:  
(A) LÄNGE: 785 Basenpaare  
(B) TYP: Nukleinsäure  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

50 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
hergestellte partielle cDNA

(iii) HYPOTHETISCH: NEIN

(iii) ANTI-SENSE: NEIN

55 (vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH  
(C) ORGAN:

5 (vii) SONSTIGE HERKUNFT:  
 (A) BIBLIOTHEK: cDNA library

5 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:19

10 ggaggagaaa aaggttagttt gaccctggta aattgtaaac agttaataat ggagtttattc 60  
 tgacatgaga aattcagtaa ttgggaccag gcgcgggtgg tcttgcctgt aatccccagc120  
 actttggag cccgaggcag gcagatcaca agttcaggag ttcgagacca gcctgaccaal80  
 15 catggtaaaa ccctgtctt actaaaaata caaaaattag ccgggggtgg tgacatgtgc240  
 ctgttaatcc cagtaactca ggaggttaagg caggagaatc gcttaaacc accaggcgga300  
 gtttgcagtg agccgagatt gcaccactgc actccagcct ggttggcaga gtgagactcg360  
 tctcaaaaaa aagaaaagaaa attagtaatt gtaagtaccc ctgataagca aattagtaat420  
 tgtcaatacc cctgttaagc aattccttt tgcaagtatat ttctgaaatg acagaatgct480  
 20 gttttaaaaa caaagaaata aaatcctgtc cctgactcgg tctaaatatt ttttaaagtc540  
 tatttttgt tttgttgc ggtactaaga ggcaatttaa aagtataaaaa ctgctttgtta600  
 tccatgaggg tttcatttgc ttttagcagc agtgagcttc tattaaatgt atatgtcatt660  
 tattttgtt aagtggctt cagcaaacct cagtcattt ctatgcagg gtattgcga720  
 acaacttgtt ttcttattaaat cgtgtttca attaaaagac cacagacttc tggaaaaaaa780  
 aaaaaa 785

(2) INFORMATION ÜBER SEQ ID NO: 21:

25 (i) SEQUENZ CHARAKTERISTIK:

- (A) LÄNGE: 901 Basenpaare
- (B) TYP: Nukleinsäure
- (C) STRANG: einzel
- (D) TOPOLOGIE: linear

30 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
 hergestellte partielle cDNA

35 (iii) HYPOTHETISCH: NEIN

35 (iii) ANTI-SENSE: NEIN

40 (vi) HERKUNFT:

- (A) ORGANISMUS: MENSCH
- (C) ORGAN:

40 (vii) SONSTIGE HERKUNFT:

- (A) BIBLIOTHEK: cDNA library

45 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:21

45 ccacgatgtt gctgctcagc ctgaccctaa gcctggttct cctcggttcc tccctggggct 60  
 gcgccattcc tgccatcaaa cccggactga gcttcagcca gaggattgtc aacggggaga120  
 atgcagtgtt gggctcctgg ccctggcagg tgcctctgca ggacagcagc ggcttccact180  
 tctcggtgg ttctctcatc agccagtcct ggggtggcac tgctgcccac tgcaatgtca240  
 50 gccctggccg ccattttgtt gtcctgggg agtatgaccg atcatcaaacc gcagagccct300  
 tgcaggttct gtccgtctct cggggcattt cacaccctag ctgaaactct accaccatga360  
 acaatgacgt gacgctgctg aagctcgctt cgccagccca gtacacaaca cgcatctcg420  
 cagtttgcct ggcattctca aacgaggctc tgactgaagg cctcacgtgt gtcaaccaccg480  
 55 gctggggctcg cctcagtgcc gtgggcaatg tgacaccagc acgtctgcag cagttggctt540  
 tgcctctgtt cactgtgaat cagtggccgc agtactgggg ctcaagtatc actgactcca600  
 tgcattgtgc aggtggcgca ggtgcctcct cgtggccaggg tgactccgga ggcctcttg660  
 tctgcccagaa gggaaacaca tgggtgctt ttggattgt ctcctggggc accaaaaact720

5 gcaatgtgcg cgcacctgct gtgtatactc gagtttagcaa gttcagcacc tggatcaacc780  
10 aggtcatagc ctacaactga gctcaccaca ggcctcccc agctcaaccc attaaagacc840  
15 caggccctgt cccatcatga aaaaaaaaaa gggaaaagggcg gcccacaact900  
g 901

5 (2) INFORMATION ÜBER SEQ ID NO: 24:

10 (i) SEQUENZ CHARAKTERISTIK:

15 (A) LÄNGE: 560 Basenpaare  
(B) TYP: Nukleinsäure  
(C) STRANG: einzel  
(D) TOPOLOGIE: linear

20 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
15 hergestellte partielle cDNA

25 (iii) HYPOTHETISCH: NEIN

30 (iii) ANTI-SENSE: NEIN

20 (vi) HERKUNFT:

25 (A) ORGANISMUS: MENSCH  
(C) ORGAN:

30 (vii) SONSTIGE HERKUNFT:

35 (A) BIBLIOTHEK: cDNA library

40 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:24

30 gggacggctc agcgacgcca cggccagca cgctcgcgtc ctccccagca acagttactc 60  
aaagctaattc agatagcgaa agaaggcaggaa gagcaagtca agaaaatacgg tgaaggagtc120  
cttcccaaag ttgtcttaggt cttccgcgc cgggtcctgg ttttcgtcgt caacaccatg180  
gacagctccc gggAACCGAC tctggggcgc ttggacgcgc ctggcttctg gcaggtctgg240  
35 cagcgcttg atgcggatga aaaagggttac atagaagaga aggaactcga tgctttctt300  
ctccacatgt ttagtggaaact gggtaactgat gacacggtca tgaaagcaaa tttgcacaag360  
gtgaaacagc agtttatgac tacccaaagat gcctctaaag atggtcgcatt tcggatgaaa420  
gagcttgctg gtatgttctt atctgaggat gaaaacttgc ttctgctctt tcggccggaa480  
40 aaccactgg acaaggcagcg tggagtttat gcagatttg cgcaaataatg acgctgacag540  
cagtggcttt atatcagctg 560

40 (2) INFORMATION ÜBER SEQ ID NO: 25:

45 (i) SEQUENZ CHARAKTERISTIK:

45 (A) LÄNGE: 565 Basenpaare  
(B) TYP: Nukleinsäure  
(C) STRANG: einzel  
(D) TOPOLOGIE: linear

50 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
hergestellte partielle cDNA

55 (iii) HYPOTHETISCH: NEIN

55 (iii) ANTI-SENSE: NEIN

## (vi) HERKUNFT:

(A) ORGANISMUS: MENSCH  
 (C) ORGAN:

## 5 (vii) SONSTIGE HERKUNFT:

(A) BIBLIOTHEK: cDNA library

## (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:25

10 gtcggccgtg cgccgcgtgc tgctggccct ggcgcgtgcc tccgtgcctt ggcgcggg 60  
 cgccgtcccc gcctccggcc acctcaagca ctggacggg acgcgcactt ggcgcaggct120  
 ctatgacaag agcgaccctt actatgagaa ctgctgggg ggcgcggagc tgcgtggaa180  
 gtcggccgtca gacctgcctt acctgcctc caactggcc aacaccgcct cctcacttgt240  
 ggtggccccc cgctgcgagc tcaccgtgtg gtcccgcaaa ggcaaggcg gcaagacgca300  
 15 caaggctctt gccggcacct acccgccctt ggaggagttt cgcgggggca tcttaggaga360  
 ctggccaac gctatctccg cgctctactg caggtgcagc tgatgcattt ctggctctc420  
 atctgcagct tccacagatg gccaagcccc tcactcagcc catccctggg ctctgctccg480  
 gggcccaag acccaggagg aggagcgttc tgcctgcccc ctcccacctc ccctgcaata540  
 cagccttgtt gcaatgttcc acgtc 565

## 20 (2) INFORMATION ÜBER SEQ ID NO: 27:

## (i) SEQUENZ CHARAKTERISTIK:

(A) LÄNGE: 553 Basenpaare  
 (B) TYP: Nukleinsäure  
 (C) STRANG: einzel  
 (D) TOPOLOGIE: linear

25 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
 30 hergestellte partielle cDNA

## (iii) HYPOTHETISCH: NEIN

## (iv) ANTI-SENSE: NEIN

35 (vi) HERKUNFT:  
 (A) ORGANISMUS: MENSCH  
 (C) ORGAN:40 (vii) SONSTIGE HERKUNFT:  
 (A) BIBLIOTHEK: cDNA library

## (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:27

45 cttgttaagaa ggctcatgcc attgaccctc ttaattctct cctgtttggc ggagtgacaa 60  
 tggcgaggc tgaaggcttg atccagatgg aagctgtgaa agtgaaaaaca ttaaaagtctt120  
 tgacggacc tccagcaatg ggcctctgtt agggcaagtc tgcagtaaaa acgactatgt180  
 tcctgtatccat gaatcatcat ccagtttcaa atagttactg actcagcaag240  
 aattcaaaaga actgttttg tcttctacta cttttctct cctaacatct ctattccaaa300  
 ctgtggcggt tacctggata ctttggagg atcccttacc agccccaaattt accccaaagcc360  
 50 gcatcctgag ctggcttatt gtgtgtggca catacaagtg gagaaagatt acaagataaa420  
 actaaacttc aaagagattt tccttagaaat agacaaacag tgcaaatttg attttcttgc480  
 catctatgat ggcccttca acaactttgg gctgattggg caagtctgtg ggcgtgtggaa540  
 ttgccaactt gga 553

## 55 (2) INFORMATION ÜBER SEQ ID NO: 28:

## (i) SEQUENZ CHARAKTERISTIK:

- (A) LÄNGE: 220 Basenpaare
- (B) TYP: Nukleinsäure
- (C) STRANG: einzeln
- (D) TOPOLOGIE: linear

5 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung hergestellte partielle cDNA

10 (iii) HYPOTHETISCH: NEIN

(iii) ANTI-SENSE: NEIN

15 (vi) HERKUNFT:

- (A) ORGANISMUS: MENSCH
- (C) ORGAN:

(vii) SONSTIGE HERKUNFT:

20 (A) BIBLIOTHEK: cDNA library

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:28

25 cgtgcctgcg ggagggcgcc gagcatgtcg ctggagtcgg ggcgcagaccc tgccttaccc 60  
gcccctcaac tggccaaaca cccggcttctc acttgggtg gccgcgtgt gcgagctcac 120  
cttgggtcc cggcaatgca atgcgggcaa gagggacaat ttctgtgggg tgagctagcg 180  
ttgactggga ggagtaccgc cgggggcatt ttaggagact 220

(2) INFORMATION ÜBER SEQ ID NO: 29:

30 (i) SEQUENZ CHARAKTERISTIK:

- (A) LÄNGE: 500 Basenpaare
- (B) TYP: Nukleinsäure
- (C) STRANG: einzeln
- (D) TOPOLOGIE: linear

35 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung hergestellte partielle cDNA

40 (iii) HYPOTHETISCH: NEIN

(iii) ANTI-SENSE: NEIN

(vi) HERKUNFT:

- (A) ORGANISMUS: MENSCH
- (C) ORGAN:

45 (vii) SONSTIGE HERKUNFT:

- (A) BIBLIOTHEK: cDNA library

50 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:29

gagccactct acagactttg atattcacat gtagcgtctg catctgaaat tggggat 60  
tctgtcccac ctgcaccctt caccggc tggtagttc ttgaggacaa ggacttcatc 120

5 attttcaaac attattggc aaataaatga agaaataggc tgcatcctt ctcttatcc180  
tttgacctcc tctatcatcc tgctgttac ttccagaagg agaagaaaca gcttcacagg240  
aaaagttagag gagattttcc catttgggt aaagtgccaa atcagaatgt gaaataaggaa300  
ttctgggctc tgtaccaggc atttactcct atgctgttag ctgatgttaa agagggtgga360  
tttctttcc cttaggtctc accttctgtg ccttcaggg aagttgggtg gaagttgaa420  
tggttgttgc ttgtcgat tggatgtat taaggagggc tgtaatggaa cgaatacaat480  
ggttattgtt gtagatgtt 500

10 (2) INFORMATION ÜBER SEQ ID NO: 30:

15 (i) SEQUENZ CHARAKTERISTIK:

- (A) LÄNGE: 298 Basenpaare
- (B) TYP: Nukleinsäure
- (C) STRANG: einzel
- (D) TOPOLOGIE: linear

20 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
hergestellte partielle cDNA

25 (iii) HYPOTHETISCH: NEIN

30 (iii) ANTI-SENSE: NEIN

(vi) HERKUNFT:

- (A) ORGANISMUS: MENSCH
- (C) ORGAN:

35 (vii) SONSTIGE HERKUNFT:

- (A) BIBLIOTHEK: cDNA library

40 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:30

45 aagaaaagca acatagtggg gtttctgtca atctgtcctc ggctgccctt ctcatttgg 60  
gatggaccc tggaaagcaag cttgttaggt gccctctgtg gctccagcc ttaccggaa120  
tgtgtgtcat gtttttaact tcagggaaac ggtatcctgt cactgggtta tggatgagc180  
atggagaaga ggcaccagcc acgattcctt cctaagcatc tcctgttctg actgctcatg240  
aattgaagaa actgaccctt gtgtcaaaa aaaaaaggca aagagggagg gtggcgga 298

50 (2) INFORMATION ÜBER SEQ ID NO: 31:

55 (i) SEQUENZ CHARAKTERISTIK:

- (A) LÄNGE: 970 Basenpaare
- (B) TYP: Nukleinsäure
- (C) STRANG: einzel
- (D) TOPOLOGIE: linear

60 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
hergestellte partielle cDNA

65 (iii) HYPOTHETISCH: NEIN

70 (iii) ANTI-SENSE: NEIN

(vi) HERKUNFT:

- (A) ORGANISMUS: MENSCH

## (C) ORGAN:

## (vii) SONSTIGE HERKUNFT:

(A) BIBLIOTHEK: cDNA library

5

## (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:31

10 ggcgtgccc agactcccaa agtgcttagga ttacacagcc ctcgcccctc gggcccttc 60  
 tgcccttgtg ggagttgcca gtcctctat cctcccagta caaagctggc ttccacttct120  
 15 gccccggctc ccacatcagc gaggactggg tggtcaccgc tgcccaccag gggtcaggtg180  
 aggcttgcgc ttgaggagtt gggcccacc aggtgtcagg gaatgacact catccccctt240  
 cccaaacccca actacacccct gcaaccaaca gaaaaagggt cggaaatgaa ctgcaggctg300  
 gggaaaccagc acccccatcc cccggccaca ggcagagccg cctctgcagg tgacacccac360  
 20 cccaggccgt gcacccccc tccacccttgc caggccaccc agacggcagc ttggggaaac420  
 ctgggagggtc ccgtaccctc actgtgcagg tggggaaatt tagaccctga aaaagggatg480  
 ccctgagatc accatgagat tgaggggcaa gcagggtca ccctgactgg ctcacttccc540  
 25 aggcacccccc atgagcccaag gcacccgcctg ccacccctcac tctccaggaa gagccaccgc600  
 gtgggtggccg ggatctgtg gtggccaggg cgtctgaccc tggctctcac ccggaggcca660  
 tccaggtgct gaggatggct aacgctaagg ccacacagcc agggagagga ggtggctcg720  
 gacaccacga tgggacacac ccacctctgg gagaggaggg tgaactccgac agcccttgcc780  
 tgccaggatg gacgctggac tctggagggc atcgtgtcct ggagcagcac cagcacctcc840  
 tggtgtcacc aggcgtggat gcccgcata taaaattcat gcctgggaag gcgtggaggc900  
 cgagaccacc cccccctaccc cgcatactgtg taaaacacaa ataaagccat tgagagcaac960  
 tgcaaaaaaa 970

25

## (2) INFORMATION ÜBER SEQ ID NO: 35:

## (i) SEQUENZ CHARAKTERISTIK:

30 (A) LÄNGE: 1032 Basenpaare  
 (B) TYP: Nukleinsäure  
 (C) STRANG: einzel  
 (D) TOPOLOGIE: linear

35 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
 hergestellte partielle cDNA

(iii) HYPOTHETISCH: NEIN

(iii) ANTI-SENSE: NEIN

40

## (vi) HERKUNFT:

(A) ORGANISMUS: MENSCH  
 (C) ORGAN:

## 45 (vii) SONSTIGE HERKUNFT:

(A) BIBLIOTHEK: cDNA library

## (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:35

50 tagccatgaa tgggataagc ggagatcagg gaactttcta gccagggaga ccaacaatga 60  
 acataaataa gtaaatgata cagttatgtg ggtgacgaat gcgggggagg aaaactaaaa 120  
 ctagaccagg tgaggtgaat cgggaacact ggaggggtgga ggtggggcag gtgggttgaa 180  
 gggtgggcct cattgacaga gtgacatttg ggcaactatt gaaggaagtg ggtggaggag 240  
 gtggcagcca ggaggtatc cagcttaggaa ggccttcca ggcagagggaa cattctagag 300  
 55 caaaggctct acagtgttgt ctggccgggg atgttgctgg aacaacaaga ctgggtgtggc 360  
 tgccccaag taagtgaagg tcaaagctgg gtagaaggct gagggggcag ggctgcgtac 420  
 acacccatcaca cgcttccatgc ctgagtgcccttggc gggcaagcag gtcacaatct 480

5 tatccaagggt tgacctcaact tttggactcc aggccttctt tgaaactga ttcaactctg 540  
 agtcaagggt ttgcttcatt caacatggc aagttaaatt ctgatatgca gcctactgct 600  
 ctctcctggg ccaattcaaa agtggctca ttacccactg ttaattacag ttccggc 660  
 aaagtttctt cctaaaagaa agacaagggg ctttgctt gactctggct cttgcactt 720  
 tccccctcac agaccagagc cccagggaa caaatctgg ttacccaagc agggactcg 780  
 gtatatttag gggcttcacc gaaaggctg ttggacagag gctgtttca gagggcttg 840  
 tgcaattgca gagacttctt tgcatcccct agtgaatagg ggaaccattt ctggctcct 900  
 ctttggacct cttggcttag accttttagt gggagttct tgaagttag gctgtggc 960  
 10 aacagcagac ctttagtagag ggggttcca ctcagcattt cacaaggcac agggcttgg 1020  
 caggttggat ct 1032

**(2) INFORMATION ÜBER SEQ ID NO: 36:**

**(i) SEQUENZ CHARAKTERISTIK:**

15 (A) LÄNGE: 1400 Basenpaare  
 (B) TYP: Nukleinsäure  
 (C) STRANG: einzel  
 (D) TOPOLOGIE: linear

20 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
 hergestellte partielle cDNA

(iii) HYPOTHETISCH: NEIN

25 (iii) ANTI-SENSE: NEIN

**(vi) HERKUNFT:**

(A) ORGANISMUS: MENSCH  
 (C) ORGAN:

30 (vii) SONSTIGE HERKUNFT:

(A) BIBLIOTHEK: cDNA library

**(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:36**

35 aagaaaagca acatagtggg gtttctgtca atctgtcctc ggctgccctt ctcatttgg 60  
 gatgggacct tgaaaacaag cttgcttaggt gcccctctgtg gctccagcct ttaccggaa 120  
 tgggtgcatttttaact tcagggaaagc ggtatctgt cactgggta tggatgagc 180  
 atggagaaga ggcaccagcc acgattcctt cctaagcatc tccgttctg actgctcatg 240  
 40 aattgaagaa actgaccctt gtgttcaactc tgcttcctt gattgttct atcatggtt 300  
 ccaaaaatttgcacatggaaac atgtcaaagg gccacttttg gagactctgc tccaggaaga 360  
 actagccccccatagccccccatatccccccac cataccacag acatgctgtg acttagagaa 420  
 ctacacaaac atccctggga cctagatgtt ggagaaatga ctaatttga tgcagaaact 480  
 ccatcaccca agggatgtacc ttcactgtt aacagtgtt tgctgtct gaaggattaa 540  
 45 gcaagaagtc ccaacagaag caatgcacca gtcctgtat gggatggaa ctgcaagcc 600  
 taggatgg gtgaagtccc tttctggaga ctttaatga gtacttcaaa gcactcgaca 660  
 cagatgcacg aaataggggca tcacacatac acaggcatca atactcacgg gcatatagat 720  
 gctgacatgt gtacactgac ttacgccctt cccacagcta cagataaggc ctgcggaa 780  
 tggcttcaga gacacatcg gaaccaagg ggaccagcag gtgcggagcc tggatctg 840  
 50 cttggaggag acgttccat gtgctgcctt gttcagatgtt ggtgtatgtt caagaaacag 900  
 aaacccacca caatttctca ggcaaaaagg gagttattta taaggacata agagcacaaa 960  
 gttccagtgc aagagataca tccaggctgc acaagctccg ggagtggggc ctggcaagcc 1020  
 aaaagaaaacc aaagttgtc ttgccttcgtt ttccttttc tgaagccaca tagcctttt 1080  
 tgactgtta tctttgcattt gttttgttt tctttagt tctctgaagc cagctttcc 1140  
 55 tggactca tcccttgatt aaatatggac attctagtt cacatcaactt cctaattcag 1200  
 ggaccaacag agactggta gcatgttgc atcccaactt ccaaatggat taggggggtt 1260  
 tgatggcg ggttggggga agggggccct ggtttggcc aggggggttgg gggccatgtg 1320  
 gggacatagg ccaaaggggc tggagcctaa tgggggcagg tttgacaac aaatgggtc 1380  
 atttcttggg acagacattc 1400

## (2) INFORMATION ÜBER SEQ ID NO: 37:

## (i) SEQUENZ CHARAKTERISTIK:

5 (A) LÄNGE: 366 Basenpaare  
 (B) TYP: Nukleinsäure  
 (C) STRANG: einzeln  
 (D) TOPOLOGIE: linear

10 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
 hergestellte partielle cDNA

(iii) HYPOTHETISCH: NEIN

15 (iii) ANTI-SENSE: NEIN

## (vi) HERKUNFT:

(A) ORGANISMUS: MENSCH  
 (C) ORGAN:

20 (vii) SONSTIGE HERKUNFT:  
 (A) BIBLIOTHEK: cDNA library

## (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO:37:

25 tgcgtgcataataaccaca atataccatc atccttctta ttatgtgga gactagttca 60  
 atcgatttt ctgtcaccta agaatttacc tacccagga gcctgccttc cacacataca120  
 ttaatagaca ccaaccagta atgtcaaaag gaaaaattac aaacccagaa aattaaagtcl80  
 attctgcact tgcccttggt ttaacaggca tttcactctt ggcaccttc ctgtcctatc240  
 30 attaataagc atcttattga tacagttat actccaaatt ctccaggctt gtgaaagttt300  
 cctcaggatt gcttgaaaat gaaagtccctg gccaggtgcg cagtggctca ggcctgtaat360  
 cccagc 366

## (2) INFORMATION ÜBER SEQ ID NO: 39:

35 (A) LÄNGE: 130 Aminosäuren  
 (B) TYP: Protein  
 (C) STRANG: einzeln  
 (D) TOPOLOGIE: linear

40 (ii) MOLEKÜLTYP: ORF

(iii) HYPOTHETISCH: ja

45 (vi) HERKUNFT:  
 (A) ORGANISMUS: MENSCH

50 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 39:

VLPRRCLVFV VNTMDSSREP TLGRILDAAGF WQVWQRFDAD EKGYIEEKEL DAFFLHMLMK 60  
 LGTDDTVMKA NLHKVKQQFM TTQDASKDGR IRMKELAGMF LSEDENFLLL FRRENPLDKQ120  
 130 RGVYADLAQI

## (2) INFORMATION ÜBER SEQ ID NO: 40:

5 (A) LÄNGE: 127 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

10 (ii) MOLEKÜLTYP: ORF

15 (iii) HYPOTHETISCH: ja

15 (vi) HERKUNFT:

15 (A) ORGANISMUS: MENSCH

15 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 40:

20 EHTSKLFHPN ATIFRGILGS HKLLFHLVQI CFHDRVISTQ FHQHVEKESI EFLLFYVTFF 60  
IRIKALPDLP EASGVQAPQS RFPGAVHGVD DEDQAPARKD LDNFGKDSFT VFLDLLSCFF120  
RYLISFE 127

## (2) INFORMATION ÜBER SEQ ID NO: 41:

25 (A) LÄNGE: 139 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

30 (ii) MOLEKÜLTYP: ORF

35 (iii) HYPOTHETISCH: ja

35 (vi) HERKUNFT:

35 (A) ORGANISMUS: MENSCH

35 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 41:

40 ELLRGRRAVA GVGRRPALPA LQLGOHRLLT CGGPALRAHR VVPARQGGQD AQVLCRHLPA 60  
PGGVPPGHLR RLVQRYLRAL LQVQLMHCWS LICSFHRVPS PSLSPSLGSA PGPQDPGGGA120  
FCLPPPTSPA IQPLCSCPR 139

## (2) INFORMATION ÜBER SEQ ID NO: 42:

45 (A) LÄNGE: 133 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

50 (ii) MOLEKÜLTYP: ORF

50 (iii) HYPOTHETISCH: ja

## (vi) HERKUNFT:

(A) ORGANISMUS: MENSCH

5

## (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 42:

10 SPLRPLLLAL ALASVPCAQG ACPASADLKH SDGTRTCAKL YDKSDPYYEN CCGGAELSLE 60  
SGADILPYLPS NWANTASSLV VAPRCELTIVW SRQKAGKTH KFSAGTYPRL EYRRGILGD120  
WSNAISALYC RCS 133

## (2) INFORMATION ÜBER SEQ ID NO: 43:

15 (A) LÄNGE: 117 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

20

(ii) MOLEKÜLTYP: ORF

(iii) HYPOTHETISCH: ja

25

## (vi) HERKUNFT:

(A) ORGANISMUS: MENSCH

## (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 43:

30

PRCARCCWPW PLPPCLAPRA PAPPPTTSST RTGRALAPSS MTRATPTMRT AAGAPSCRWS 60  
RAQTCPTCPP TGPTPPPHLW WPRAASSPCG PGKARRARRT SSLPAPTRAW RSTAGAS 117

## (2) INFORMATION ÜBER SEQ ID NO: 44:

35

(A) LÄNGE: 160 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

40

(ii) MOLEKÜLTYP: ORF

(iii) HYPOTHETISCH: ja

45

## (vi) HERKUNFT:

(A) ORGANISMUS: MENSCH

50

## (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 44:

RLDPDGSCES ENIKVFDGTS SNGPLLGQVC SKNDYVPVFE SSSSTLTQI VTDSARIQRT 60  
VFVFYYFFSP NISIPNCGGY LDTLEGSFTS PNYPKPHPEL AYCVWHIQVE KDYKIKLNFK120  
EIFLEIDKQC KFDFLAIYDG PFNNFGLIGQ VCGRVDCQLG 160

## (2) INFORMATION ÜBER SEQ ID NO: 45:

5 (A) LÄNGE: 73 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

10 (ii) MOLEKÜLTYP: ORF

15 (iii) HYPOTHETISCH: ja

15 (vi) HERKUNFT:

15 (A) ORGANISMUS: MENSCH

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 45:

20 RACGRAPSMS LESGADRALP ALQLGQHRL TCGGRVLRAH LVVPAMQCGQ EGQFLWGELA60  
LTGRSTAGGM LGD 73

## (2) INFORMATION ÜBER SEQ ID NO: 46:

25 (A) LÄNGE: 78 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

30 (ii) MOLEKÜLTYP: ORF

(iii) HYPOTHETISCH: ja

35 (vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 46:

40 RLHLKLFLHL SHLHPSPQAV SFLRTRTSSF SNIIGQINEE IGCILSLYPL TSSIILLSS60  
RRRRNSFTGK VEEIFPFW 78

## (2) INFORMATION ÜBER SEQ ID NO: 47:

45 (A) LÄNGE: 50 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

50 (ii) MOLEKÜLTYP: ORF

(iii) HYPOTHETISCH: ja

5 (vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH

10 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 47:

WDLESKLARC PLWLQPLPEV WCMFLTSGKR YPVTGVWDEH GEEAPATIPS

50

15 (2) INFORMATION ÜBER SEQ ID NO: 48:

(A) LÄNGE: 70 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

20 (ii) MOLEKÜLTYP: ORF

25 (iii) HYPOTHETISCH: ja

(vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH

30 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 48:

EGIVAGASSP CSSHTPVGY RFPEVKNMHH TSGKGWSHRG HLASLLSRSH QQMRRAAEDR60  
LTETPLCCFS

70

35 (2) INFORMATION ÜBER SEQ ID NO: 49:

(A) LÄNGE: 51 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

40 (ii) MOLEKÜLTYP: ORF

45 (iii) HYPOTHETISCH: ja

(vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH

50 :

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 49:

RHPSSLPFFF EHKGQFLQFM SSQNRRCLGR NRGWCLFSML IPYPSDRIPL P

51

55 (2) INFORMATION ÜBER SEQ ID NO: 50:

(A) LÄNGE: 161 Aminosäuren  
(B) TYP: Protein

(C) STRANG: einzel  
(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: ORF

5

(iii) HYPOTHETISCH: ja

(vi) HERKUNFT:

10 (A) ORGANISMUS: MENSCH

15 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 50:

ARAQTPKVLG LHSPPRPGPF CPCGSCQSSY PPSTKLASTS AGATTSARTG WSPLPTRGQV 60  
RLELEELGPT RCQGMLIPL PNPNYTLQPT GKGSEMCRL GNQHPPAT GRAASAGDTH120  
PRPCTPPPPL QATQTAAWGN LGGPVPSLCR WGNLDPEKGM P 161

(2) INFORMATION ÜBER SEQ ID NO: 51:

20

(A) LÄNGE: 107 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzel  
(D) TOPOLOGIE: linear

25

(ii) MOLEKÜLTYP: ORF

(iii) HYPOTHETISCH: ja

30

(vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH

35

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 51:

FFCSCSQWLY LCFTQMRGRG GGLGLHAFPG MNFMMRASTP GDNRRCWCCS RTRCPPESRL 60  
HPGRQGLSES PSSPRGGCVP SWCHEPPLP GCVALALAIL STWMASG 107

40

(2) INFORMATION ÜBER SEQ ID NO: 52:

45

(A) LÄNGE: 118 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzel  
(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: ORF

(iii) HYPOTHETISCH: ja

50

(vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH

## (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 52:

5           QQEVLVLLQD TMPSRVQAPS WQARAVGVTL LSQRWVCIV VSRATSSPWL CGLSVSHPQH 60  
LDGLRVRAKV RRPGHHTIPA TTRWLFLESE GGRRCLGSWG CLGSEPVRS PACPSISW 118

## (2) INFORMATION ÜBER SEQ ID NO: 56:

10           (A) LÄNGE: 76 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzel  
(D) TOPOLOGIE: linear

15           (ii) MOLEKÜLTYP: ORF

15           (iii) HYPOTHETISCH: ja

20           (vi) HERKUNFT:

20           (A) ORGANISMUS: MENSCH

## (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 56:

25           VPRPWWASRS QSYPRLTSLL DSRPSLETDS TLSQGFASFN MVKLNSDMQP TALSWANSKV60  
ASLPTVNYSF PGKVSS 76

## (2) INFORMATION ÜBER SEQ ID NO: 57:

30           (A) LÄNGE: 78 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzel  
(D) TOPOLOGIE: linear

35           (ii) MOLEKÜLTYP: ORF

35           (iii) HYPOTHETISCH: ja

40           (vi) HERKUNFT:

40           (A) ORGANISMUS: MENSCH

## (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 57:

45           RSNLPKPCAL CNAEWNPPLL RSAVDQQPHF QETPTKRSKP RGPKRRPAMV PLFTRGCKEV60  
SAIAQAPLKT ASVQQPFR 78

## (2) INFORMATION ÜBER SEQ ID NO: 58:

50           (A) LÄNGE: 136 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzel  
(D) TOPOLOGIE: linear

5 (ii) MOLEKÜLTYP: ORF

(iii) HYPOTHETISCH: ja

10 (vi) HERKUNFT:

(A) ORGANISMUS: MENSCH

15 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 58:

ALRQEACEVC TQPCPLSLLP SFDLHLLGGS HTSLVVPATS PARPHCRAFA LECSSAWKAL 60  
PSWIASWLPP PPPTSFNSCP NVTLSMRPTL QPTCPTSTLQ CSRFTSPGLV LVFLPRIRHL120  
16 TYCIIYLFMF IVGLPG 136

20 (2) INFORMATION ÜBER SEQ ID NO: 59:

(A) LÄNGE: 115 Aminosäuren

(B) TYP: Protein

(C) STRANG: einzel

(D) TOPOLOGIE: linear

25 (ii) MOLEKÜLTYP: ORF

(iii) HYPOTHETISCH: ja

30 (vi) HERKUNFT:

(A) ORGANISMUS: MENSCH

35 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 59:

EMGEVPFWRP LMSTSKHSTQ MQQIGHHTYT GINTHGHIDA DMCTLTYALP TATDKASQSW 60  
PQRHIRNQGG PAGAEPVYLL GGDVPMCCLV QRWCSCKKQK PTTISQAKRE LIIRT 115

40 (2) INFORMATION ÜBER SEQ ID NO: 60:

(A) LÄNGE: 70 Aminosäuren

(B) TYP: Protein

(C) STRANG: einzel

(D) TOPOLOGIE: linear

45 (ii) MOLEKÜLTYP: ORF

(iii) HYPOTHETISCH: ja

50 (vi) HERKUNFT:

(A) ORGANISMUS: MENSCH

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 60:

EGIVAGASSP CSSHTPVTGY RFPEVKNMHH TSGKGWSHRG HLASLLSRSH QQMRRAAEDR60  
LTETPLCCFS 70

5 (2) INFORMATION ÜBER SEQ ID NO: 61:

10 (A) LÄNGE: 95 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: ORF

15 (iii) HYPOTHETISCH: ja

20 (vi) HERKUNFT:

(A) ORGANISMUS: MENSCH

25 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 61:

LPFCLRNCGG FLFLATTPSL NKAHHWNVSS KQIHLGTCW STLVPDVSLR PTLRGLICSC60  
GKGVSQCTHV SIYMPVSINA CVCVMPYLLH LCRVL 95

25 (2) INFORMATION ÜBER SEQ ID NO: 62:

30 (A) LÄNGE: 62 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

35 (ii) MOLEKÜLTYP: ORF

35 (iii) HYPOTHETISCH: ja

40 (vi) HERKUNFT:

(A) ORGANISMUS: MENSCH

40 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 62:

TPTSNVKRKN YKPRKLKSFC TCPWFNRHFT LGTFPVLSLI SILLIQFILQ ILQACESFLR60  
45 IA 62

(2) INFORMATION ÜBER SEQ ID NO: 63:

50 (A) LÄNGE: 69 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

(ii) MOLEKÜLTYP: ORF

(iii) HYPOTHETISCH: ja

5 (vi) HERKUNFT:  
 (A) ORGANISMUS: MENSCH

10 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 63:

15 AGITGLSHCA PGQDFHFQAI LRKLSQAWRI WSINCINKML INDRGKVPR VKCLLNQGQV60  
 QNDFNFLGL 69

(2) INFORMATION ÜBER SEQ ID NO: 67:

15 (i) SEQUENZ CHARAKTERISTIK:  
 (A) LÄNGE: 1850 Basenpaare  
 (B) TYP: Nukleinsäure  
 (C) STRANG: einzel  
 20 (D) TOPOLOGIE: linear

25 (ii) MOLEKÜLTYP: aus einzelnen ESTs durch Assemblierung und Editierung  
 hergestellte partielle cDNA

25 (iii) HYPOTHETISCH: NEIN  
 (iii) ANTI-SENSE: NEIN

30 (vi) HERKUNFT:  
 (A) ORGANISMUS: MENSCH  
 (C) ORGAN:

35 (vii) SONSTIGE HERKUNFT:  
 (A) BIBLIOTHEK: cDNA library

35 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO: 67:

40 GAGAGCCAAT TGACGTGGTA TTTATTTCCA TCGTCTTGTA ATGAAACACA CCGACACTT 60  
 TTTTGCAAGG CCTGTGTGTC ACTTTCAGGC CCTTCCTCAT TTATTCATAA TATGTGTAGC 120  
 CACGAGGTAG CATTCCCTT GGATCCTGG AGAAACATGT CTCACTTAGG GTGGACCTG 180  
 TTAGTTAATA GTTCTGCAGC TTCTGGTATT TGTAGTCTGC CCGTTGATT ACAAAATGCC 240  
 TCACTGTGAT TGTCGCTACA GTCACCACAT TCAGAGCTAG AACCAGAAG GAAAACAGAT 300  
 GCACACTGTT GAAAGGCTGG TTTGGAGTTT CTTCCGCATG TGTTTCATGC TGAAATCCTG 360  
 AATTGCCACT TGCACCTCGA TCCCTTTCA GACGAATGGG TCCTATGATG GAATCTGTTT 420  
 45 TCCATTATA TGAAGAAATG TCTCGTTGC TTCTGGAGAC ACAACCTTGA TTGCAGCGAG 480  
 ACTGGTGGTC ACTGCTATCA CATATAAAAA CTTTACACTG CAGATACACA GAGCTCATAC 540  
 TTCTCAAGAA TTTAAAGGCA TTTAAACTGGA ATCTCCCATA GTGTCCAAT AAGGGATACA 600  
 CCTTACAAGT TTCACTCTGA CTACATCCAC TCTTGATTAG GTCTAGGTT GGAGATGCAA 660  
 AGTCAGAGGT GGGAGAGGCT CTACAGGTAT CAAGAAACAC CACCAAATTG GGATCTGAGG 720  
 50 TGTGCAGACT AACTTGAAACA AAAAGAGTTT GGTTCAAATC CACATAATAT GGTGATTCAA 780  
 GTATAGTCTT TTCAAATGAA TTGGATTCAA AAAGAGCCAT GCTGGTGTAA TATTGCCCCA 840  
 GTGCATTGGT ACCTTGATT ACATCATCTT CTGTTATGTA TATTATCTCC ACTGTAGAAAT 900  
 TATGTCCCCTT TTCACACTTC ACAATAATCT GGAGTTGTTT CTGACGGGTG ATCACTTCAG 960  
 AAGTTGAGGA TGCAGAAAAG GTGATTATAT TGTTGTAAGT AATTGACTGA TCTTCTACCT1020  
 55 TTCTGATTGT ACCACATCCA TTAAGAGGGA CAGAAAATTC CACAAACATTG GATAATTG1080  
 GTCTGCAAGT TGGGTCTTT AGTTGCAAGT TATTCCCATT AGAGTTAAAAA GCCTCTAGGT1140  
 AGGATTGCT TATAATAACT CTCATCCTGT CAGAAGAGCA AGTTAAAGAT GTAGTGTGA1200

TGTTTCTGC	ATAAATTGAG	GTGTAGGAAG	CAGAAAATCC	CCGGTAAGAA	TTGGCATATA1260	
CTGTAGACAA	CACGACAGTC	AGAGAGTTG	ATGACGATT	GAAGGGTGGGA	GTCACACGGC1320	
CACAGACTTG	TCCAAATCAGG	CCAGAGTTGG	TGGAGGGGCC	ATCATAGATG	GCAAGAAAAT1380	
CAAATTGCA	CTGTTTGTCT	ATTTCTAGGA	AAATCTCTT	GAAGTTAGT	TTTATCTTGT1440	
5	AATCTTCTC	CACTGTATG	TGCCACACAC	ATAAGCCAG	CTCAGGATGC	GGCTTGGGT1500
AATTGGGGCT	GGTGAAGGAT	CCTTCCAAGG	TATCCAGGT	ACCGCCACAG	TTTGGAAATAG1560	
AGATGTTAGG	AGAGAAGAAG	TAGTAGAAGA	CAAAGACAGT	TCTTGAATT	CTTGTGAGT1620	
CAGTAACAT	TTGAAACGTC	AATGTACTGG	ATGATGATT	AAATACAGGA	ACATAGTCGT1680	
TTTACTGCA	GACTTGCCT	AGCAGAGGCC	CATTGCTGGA	GGTTCCGTCA	AAGACTTTAA1740	
10	TGTTTCACT	TTCACAGCTT	CCATCTGGAT	CAAGCCTTC	GCCTCCGCCA	TTGTCACTCC1800
	GCCAAACAGG	AGAGAATTAA	GAGGGTCAAT	GGCATGAGCC	TTCTTACAAG	1850

(2) INFORMATION ÜBER SEQ ID NO: 68:

15 (A) LÄNGE: 74 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

## 20 (ii) MOLEKÜLTYP: ORF

(iii) HYPOTHETISCH: ja

25 (vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 68:

30 STPAVHHSCP PPPSHWHEAS LAGQHCDRRR LPRQRDPAP ATSKCEPPTR KDYHNMFVRG60  
S1GARTIVGSD SCN 74

(2) INFORMATION ÜBER SEQ ID NO: 69:

35 (A) LÄNGE: 63 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

40 [View Article Online](#)

### (iii) HYDROLYTICALLY-DEGRADABLE

45 (vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH

50 (xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 69:

SNQRPRSTTP AHRPPATGTK HLWLGNTATG ADYPDSAGLM LRRPASVNLL PGKITTCSV60  
63

55 (2) INFORMATION ÜBER SEQ ID NO: 70:

(A) LÄNGE: 530 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

5

(ii) MOLEKÜLTYP: ORF

10

(vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH

15

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 70:

RLDPDGSCES ENIKVFDGTS SNGPLLGQVC SKNDYVPVFE SSSSTLTFQI VTDSARIQRT 60  
VFVFYYFFSP NISIPNCCGY LDTLEGSFTS PNYPKPHPEL AYCWHIQVE KDYKIKLNFK120  
EIFLEIDKQC KFDFLAIYDG PSTNSGLIGQ VCGRVPTFE SSSNSLTVV1 STDYANSYRG180  
20 FSASYTSIYA ENINTTSLTC SSDRMRVIIS KSYLEAFNSN GNNLQLKDPT CRPKLSNVVE240  
FSVPLNGCGT IRKVEDQOSIT YTNIITFSAS STSEVITRQK QLQIIVKCEM GHNSTVEIIY300  
ITEDDVIQSQ NALGKYNTSM ALFESNSFEK TILESPTYVD LNQTLFVQVS LHTSDPNLVV360  
FLDTCRASPT SDFASPTYDL IKSGCSRDET CKVYPLFGHY GRFQFNNAFKF LRSMSSVYLQ420  
25 CKVLICDSSD HQSRCNQGCV SRSKRDISSY KWKTDSIIGP IRLKRDRSAS GNSGFQHETH480  
AEETPNQPFN SVHLFSFMVL ALNVVTVATI TVRHVNQRA DYKYQKLQNY 530

(2) INFORMATION ÜBER SEQ ID NO: 71:

30

(A) LÄNGE: 82 Aminosäuren  
(B) TYP: Protein  
(C) STRANG: einzeln  
(D) TOPOLOGIE: linear

35

(ii) MOLEKÜLTYP: ORF  
(iii) HYPOTHETISCH: ja

40

(vi) HERKUNFT:  
(A) ORGANISMUS: MENSCH

(xi) SEQUENZ-BESCHREIBUNG: SEQ ID NO 71:

45 ESQLTWYLFP SSCNETHRHF FLQACVSLSG PSSFIHNMCs HEVAFPLASW RNMSHLGLDL60  
LVNSSAASGI CSLPVDLQNA SL 82

**Patentansprüche**

1. Eine Nukleinsäure-Sequenz, die ein Genprodukt oder ein Teil davon kodiert,  
5 umfassend
  - a) eine Nukleinsäure-Sequenz, ausgewählt aus der Gruppe Seq. ID No 14, 24, 25, 27-  
31, 35-37, 67.
  - 10 b) eine allelische Variation der unter a) genannten Nukleinsäure-Sequenzen  
oder
  - 15 c) eine Nukleinsäure-Sequenz, die komplementär zu den unter a) oder b) genannten  
Nukleinsäure-Sequenzen ist.
2. Eine Nukleinsäure-Sequenz gemäß einer der Sequenzen Seq. ID No 14, 24, 25, 27-  
20 31, 35-37, 67, oder eine komplementäre oder allelische Variante davon.
- 25 3. Nukleinsäure-Sequenz Seq. ID No 2-37, 67, dadurch gekennzeichnet, daß sie in  
Pankreasnormalgewebe erhöht exprimiert sind.
4. BAC, PAC und Cosmid-Klone, enthaltend funktionelle Gene und ihre chromosomale  
Lokalisation, entsprechend den Sequenzen Seq. ID No 2-37, 67, zur Verwendung  
als Vehikel zum Gentransfer.
- 30 5. Eine Nukleinsäure-Sequenz gemäß den Ansprüchen 1 bis 4, dadurch  
gekennzeichnet, daß sie eine 90%ige Homologie zu einer humanen Nukleinsäure-  
Sequenz aufweist.
- 35 6. Eine Nukleinsäure-Sequenz gemäß den Ansprüchen 1 bis 4, dadurch  
gekennzeichnet, daß sie eine 95%ige Homologie zu einer humanen Nukleinsäure-  
Sequenz aufweist.
- 40 7. Eine Nukleinsäure-Sequenz, umfassend einen Teil der in den Ansprüchen 1 bis 6  
genannten Nukleinsäure-Sequenzen, in solch einer ausreichenden Größe, daß sie  
mit den Sequenzen gemäß den Ansprüchen 1 bis 6 hybridisieren.
- 45 8. Ein Nukleinsäure-Sequenz gemäß den Ansprüchen 1 bis 7, dadurch  
gekennzeichnet, daß die Größe des Fragments eine Länge von mindestens 50 bis  
4500 bp aufweist.

9. Eine Nukleinsäure-Sequenz gemäß den Ansprüchen 1 bis 7, dadurch gekennzeichnet, daß die Größe des Fragments eine Länge von mindestens 50 bis 4000 bp aufweist.

5 10. Eine Nukleinsäure-Sequenz gemäß einem der Ansprüche 1 bis 9, die mindestens eine Teilsequenz eines biologisch aktiven Polypeptids kodiert.

10 11. Eine Expressionskassette, umfassend ein Nukleinsäure-Fragment oder eine Sequenz gemäß einem der Ansprüche 1 bis 9, zusammen mit mindestens einer Kontroll- oder regulatorischen Sequenz.

15 12. Eine Expressionskassette, umfassend ein Nukleinsäure-Fragment oder eine Sequenz gemäß Anspruch 11, worin die Kontroll- oder regulatorische Sequenz ein geeigneter Promotor ist.

20 13. Eine Expressionskassette gemäß einem der Ansprüche 11 und 12, dadurch gekennzeichnet, daß die auf der Kassette befindlichen DNA-Sequenzen ein Fusionsprotein kodieren, das ein bekanntes Protein und ein biologisch aktives Polypeptid-Fragment umfaßt.

25 14. Verwendung der Nukleinsäure-Sequenzen gemäß den Ansprüchen 1 bis 10 zur Herstellung von Vollängen-Genen.

30 15. Ein DNA-Fragment, umfassend ein Gen, das aus der Verwendung gemäß Anspruch 14 erhältlich ist.

35 16. Wirtszelle, enthaltend als heterologen Teil ihrer exprimierbaren genetischen Information ein Nukleinsäure-Fragment gemäß einem der Ansprüche 1 bis 10.

17. Wirtszelle gemäß Anspruch 16, dadurch gekennzeichnet, daß es ein prokaryontisches oder eukaryontische Zellsystem ist.

40 18. Wirtszelle gemäß einem der Ansprüche 16 oder 17, dadurch gekennzeichnet, daß das prokaryontische Zellsystem *E. coli* und das eukaryontische Zellsystem ein tierisches, humanes oder Hefe-Zellsystem ist.

45 19. Ein Verfahren zur Herstellung eines Polypeptids oder eines Fragments, dadurch gekennzeichnet, daß die Wirtszellen gemäß den Ansprüchen 16 bis 18 kultiviert werden.

20. Ein Antikörper, der gegen ein Polypeptid oder ein Fragment gerichtet ist, welches von den Nukleinsäuren der Sequenzen Seq. ID No 2-37, 67 kodiert wird, das gemäß Anspruch 19 erhältlich ist.
- 5 21. Ein Antikörper gemäß Anspruch 20, dadurch gekennzeichnet, daß er monoklonal ist.
- 10 22. Ein Antikörper gemäß Anspruch 20 dadurch gekennzeichnet, daß er ein Phage-Display-Antikörper ist.
- 15 23. Polypeptid-Teilsequenzen, gemäß den Sequenzen Seq. ID No 39-63, 68-71.
- 20 24. Polypeptid-Teilsequenzen gemäß Anspruch 23, mit mindestens 80%iger Homologie zu diesen Sequenzen.
- 25 25. Ein aus einem Phage-Display hervorgegangenen Polypeptid, welches an die Polypeptid-Teilsequenzen gemäß Anspruch 23 binden kann.
26. Polypeptid-Teilsequenzen gemäß Anspruch 23, mit mindestens 90%iger Homologie zu diesen Sequenzen.
- 30 27. Verwendung der Polypeptid-Teilsequenzen gemäß den Sequenzen Seq. ID No 39-63, 68-71, als Tools zum Auffinden von Wirkstoffen gegen den Pankreastumor.
- 35 28. Verwendung der Nukleinsäure-Sequenzen gemäß den Sequenzen Seq. ID No 2-37, 67 zur Expression von Polypeptiden, die als Tools zum Auffinden von Wirkstoffen gegen den Pankreastumor verwendet werden können.
- 40 29. Verwendung der Nukleinsäure-Sequenzen Seq. ID No 2-37, 67 in sense oder antisense Form.
- 45 30. Verwendung der Polypeptid-Teilsequenzen Seq. ID No 39-63, 68-71 als Arzneimittel in der Gentherapie zur Behandlung den Pankreastumor.
31. Verwendung der Polypeptid-Teilsequenzen Seq. ID No 39-63, 68-71, zur Herstellung eines Arzneimittels zur Behandlung gegen den Pankreastumor.
- 45 32. Arzneimittel, enthaltend mindestens eine Polypeptid-Teilsequenz Seq. ID No 39-63, 68-71.

33. Eine Nukleinsäure-Sequenz gemäß den Ansprüchen 1 bis 10, dadurch gekennzeichnet, daß es eine genomische Sequenz ist.

5 34. Eine Nukleinsäure-Sequenz gemäß den Ansprüchen 1 bis 10, dadurch gekennzeichnet, daß es eine mRNA-Sequenz ist.

10 35. Genomische Gene, ihre Promotoren, Enhancer, Silencer, Exonstruktur, Intronstruktur und deren Spleißvarianten, erhältlich aus den cDNAs der Sequenzen Seq. ID No 2-37, 67.

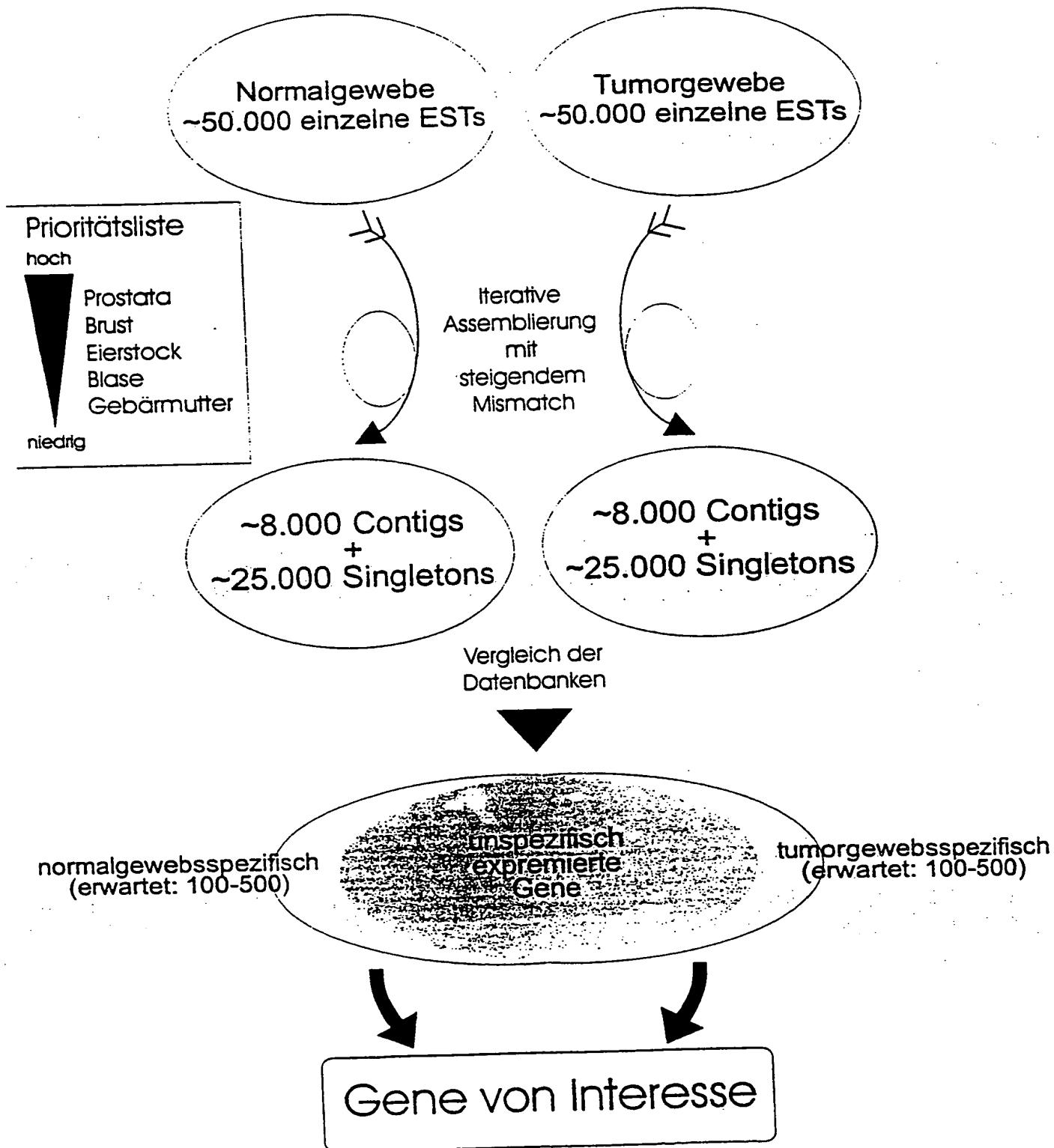
15 36. Verwendung der genomischen Gene gemäß Anspruch 33, zusammen mit geeigneten regulativen Elementen.

20 37. Verwendung gemäß Anspruch 36, dadurch gekennzeichnet, daß das regulative Element ein geeigneter Promotor und/ oder Enhancer ist.

25 38. Eine Nukleinsäure-Sequenz gemäß den Ansprüchen 1 bis 7, dadurch gekennzeichnet, daß die Größe des Fragments eine Länge von mindestens 300 bis 3500 bp aufweist.

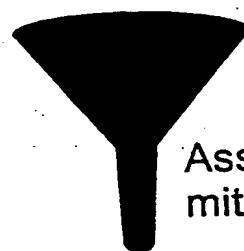
1/10

# Systematische Gen-Suche in der Incyte LifeSeq Datenbank



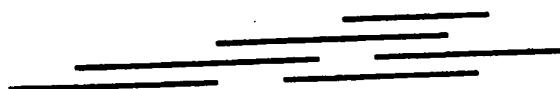
# Prinzip der EST-Assemblierung

~50.000 ESTs pro Gewebe



Assemblierung bei 0% Mismatch  
mit GAP4 (Staden)

Contigs



Singletons



In Anzahl und Länge  
zunehmende Contigs



Iterative Assemblierung  
mit steigendem  
Mismatch (1%, 2%, 4%)

5000-6000 Contigs    ~25.000 übrige Singletons



~30.000 Konsensus-  
sequenz n pro Gewebe

**metaGen**  
Computer for Comprehension

Fig. 2a

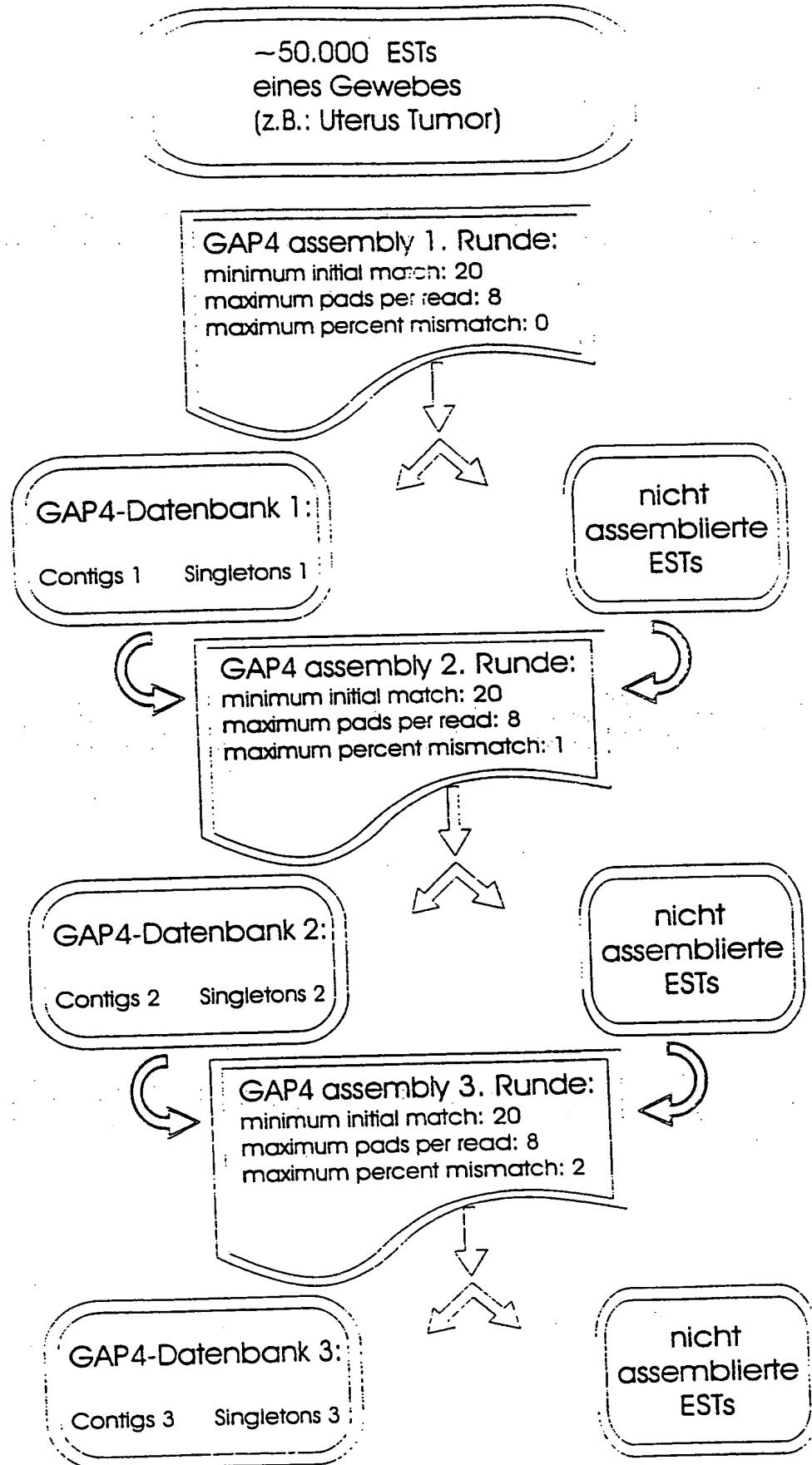


Fig. 2b1

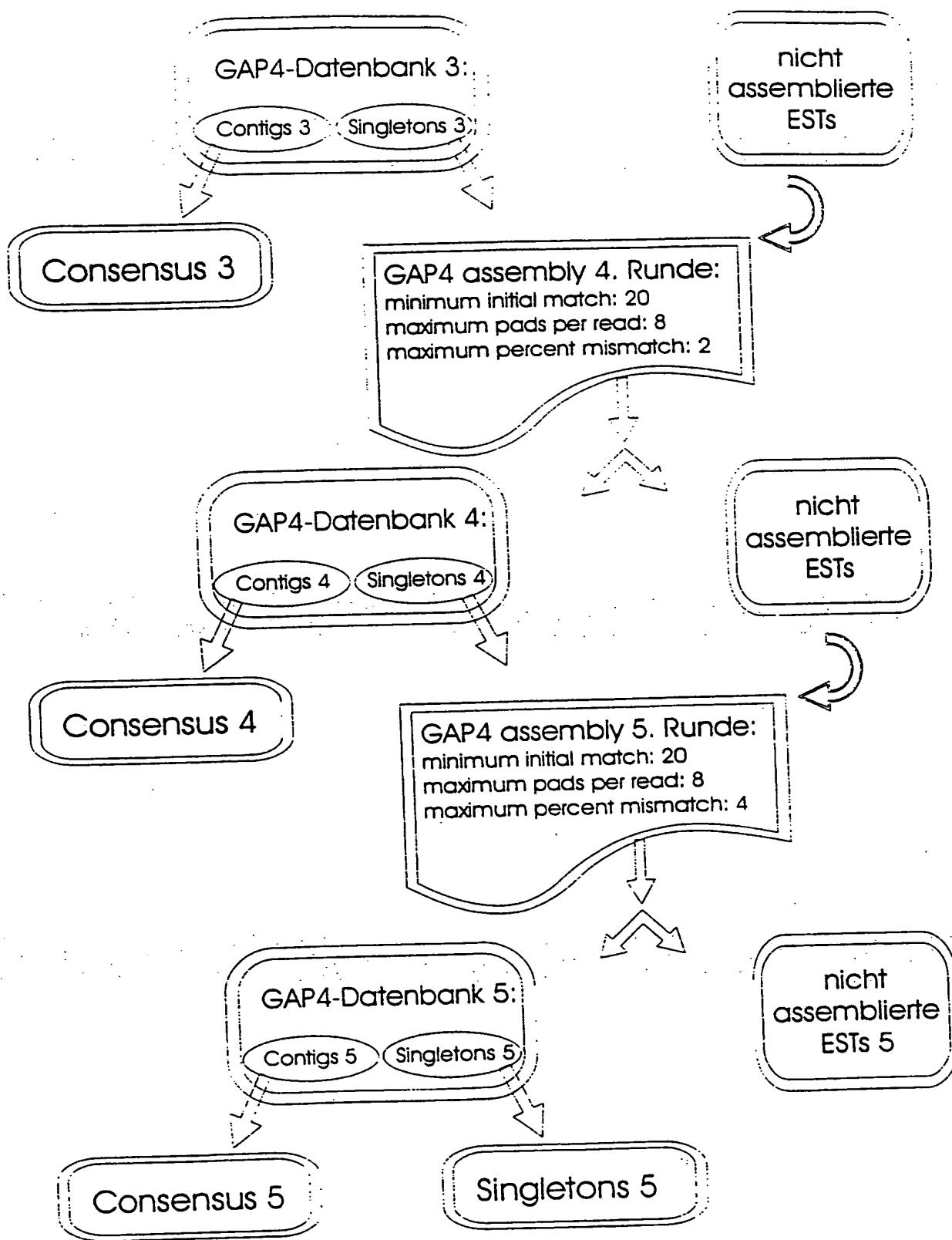


Fig. 2b2

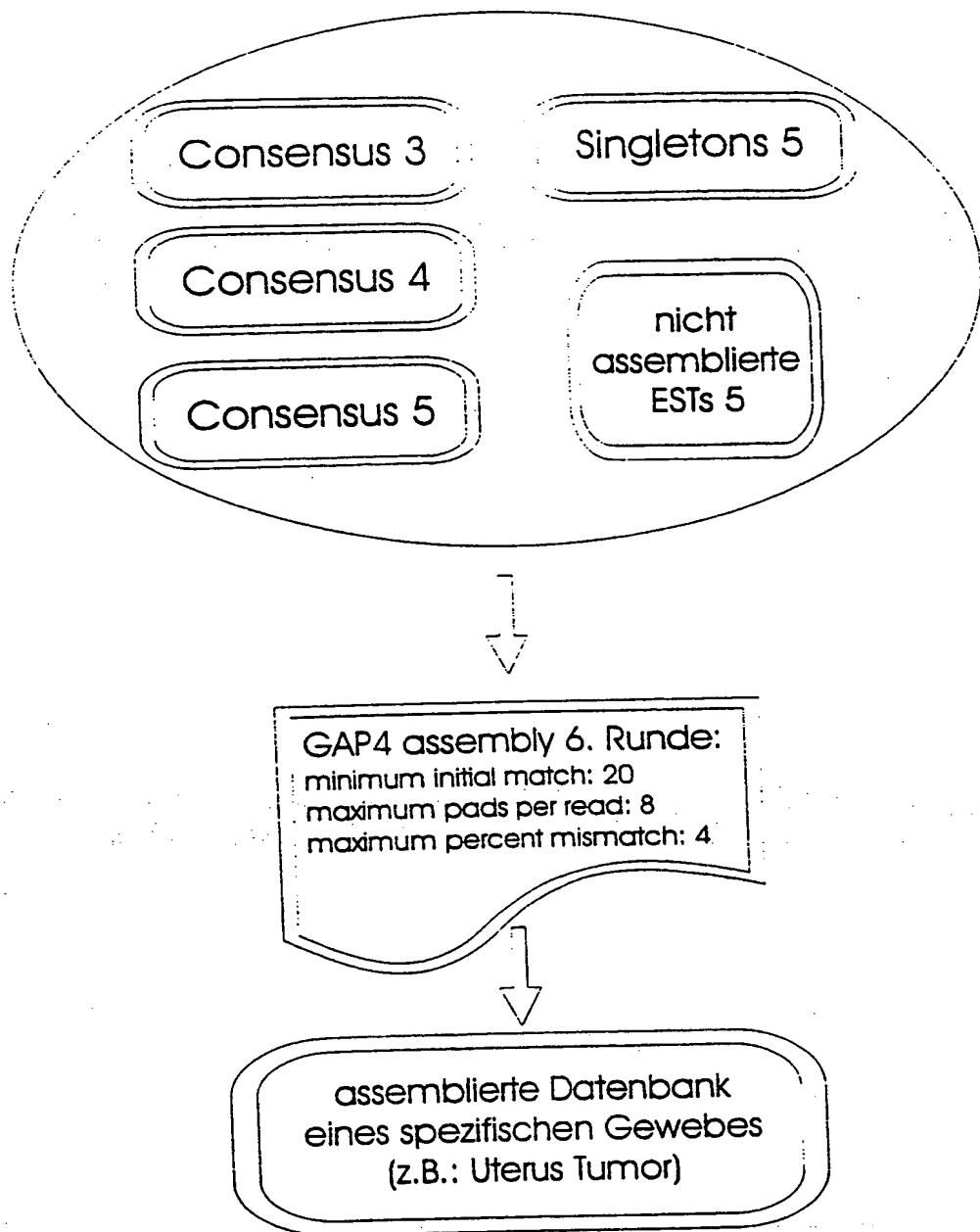


Fig. 2b3

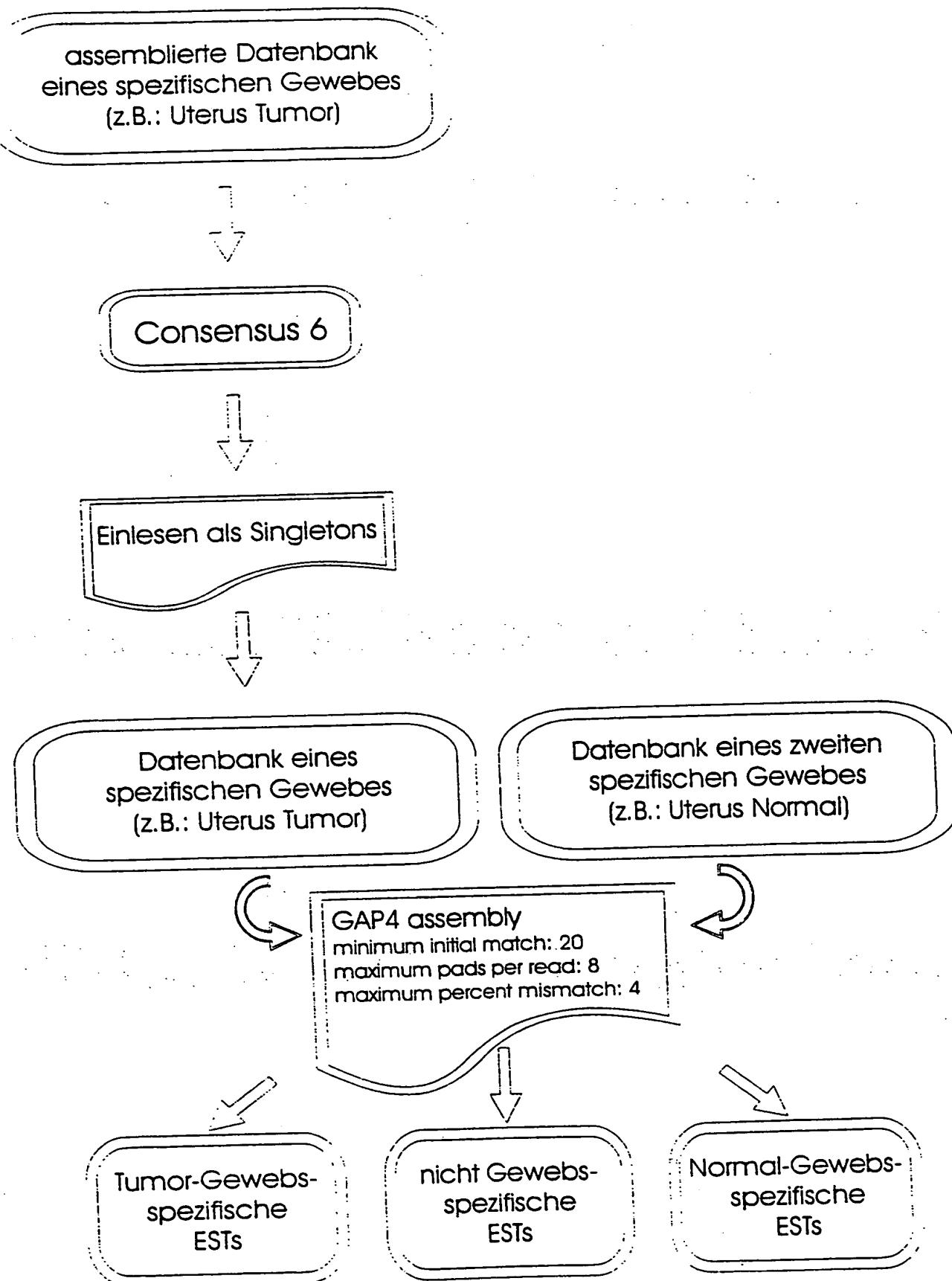


Fig. 2b4

## In silico Subtraktion der Genexpression in verschiedenen Geweben

~30.000 Konsensussequenzen  
Normalgewebe

~30.000 Konsensussequenzen  
Tumorgewebe

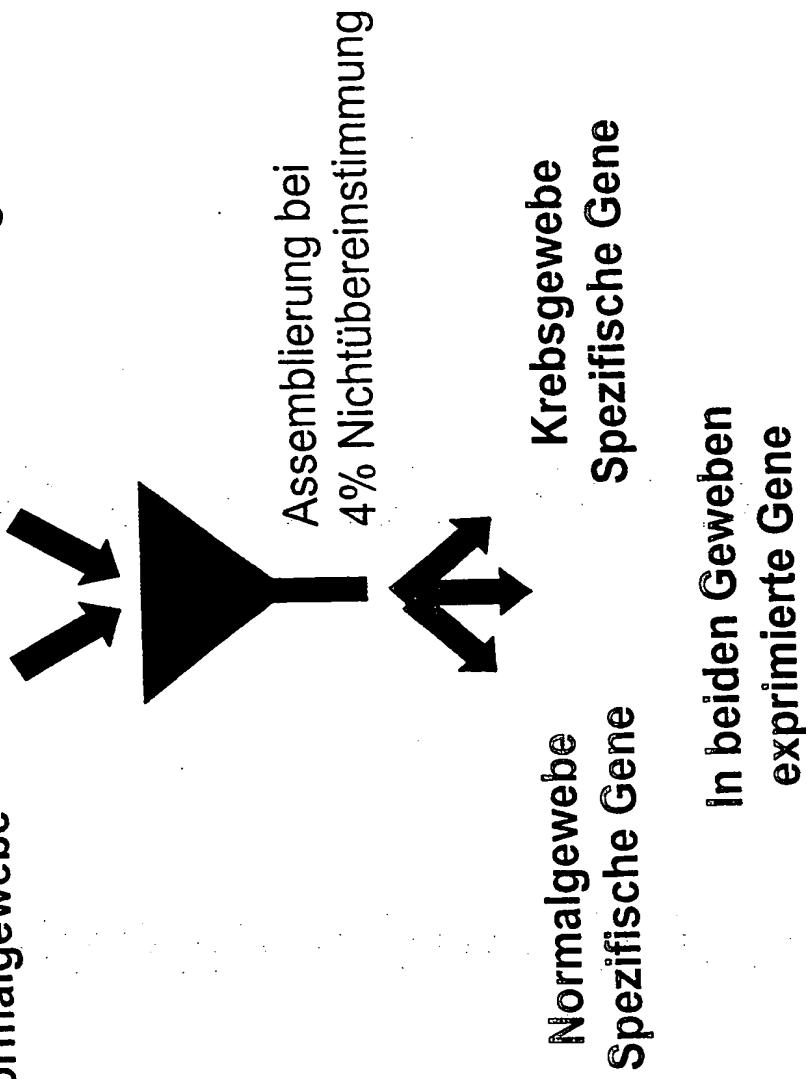


Fig. 3

8/10



Gene von Interesse

Bestimmung der  
gewebsspezifischen  
Expression über  
elektronischen Northern  
(INCYTE LifeSeq und  
öffentliche EST  
Datenbanken)

Kandidatengene für  
Tumorsuppressoren oder  
Tumoraktivatoren

. Fig. 4a

9/10

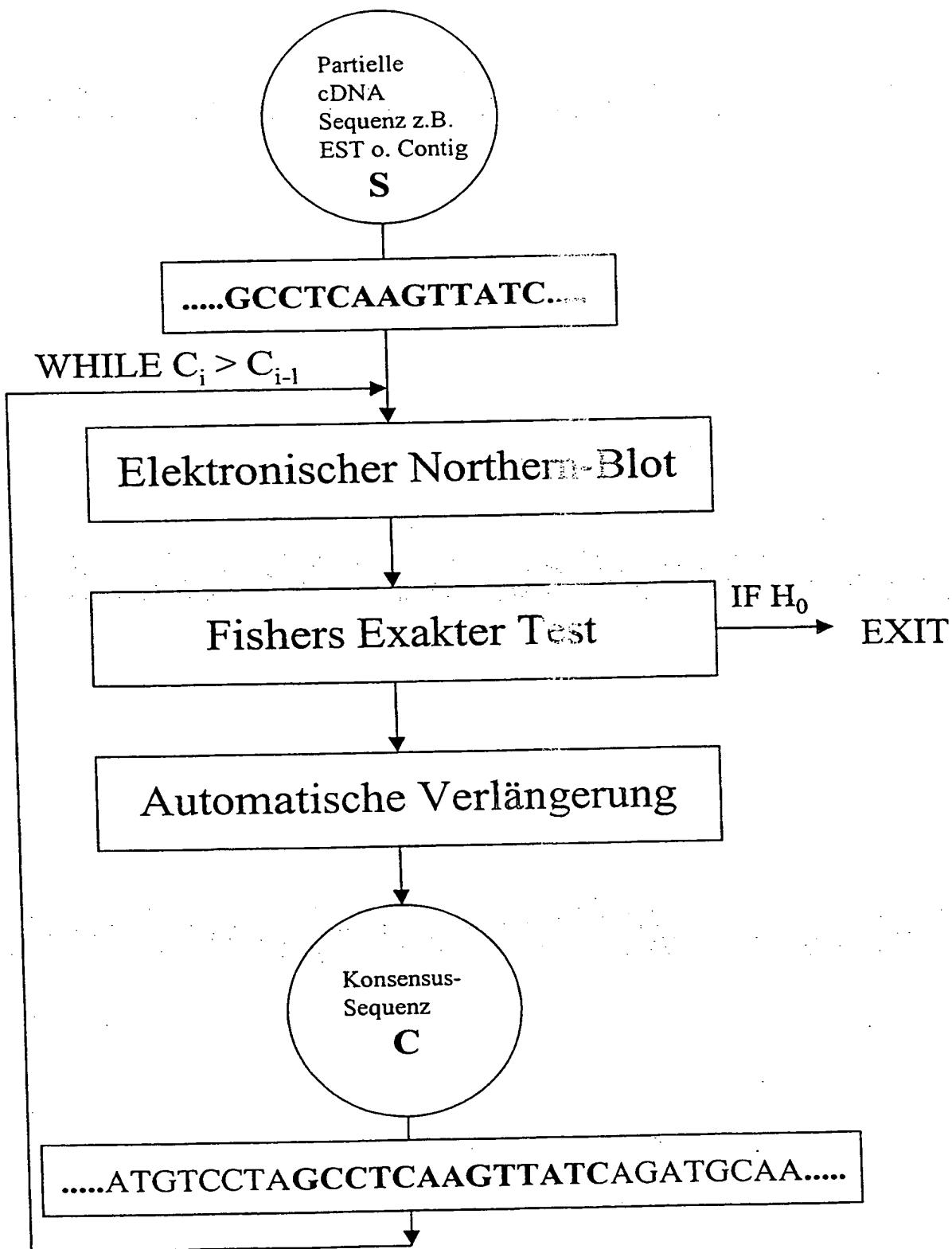


Fig. 4b

10/10

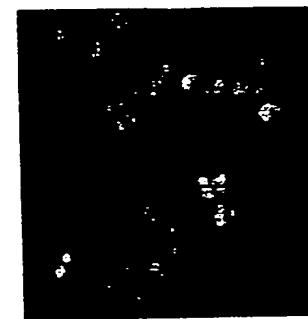
Isolieren von genomischen BAC und PAC Klonen



Chromosomal Klon-Lokalisation über FISH



Hybridisierungssignal



Sequenzierung von Klonen, die in Regionen lokalisiert sind, die chromosomal Deletionen in Prostata- und Brustkrebs aufweisen, führt zur Identifizierung von Kandidatengenen



Bestätigung der Kandidatengene durch Screening von Mutationen und/oder Deletionen in Krebsgeweben

Fig. 5

ERSATZBLATT (REGEL 26)



**PCT**

WELTORGANISATION FÜR GEISTIGES EIGENTUM

Internationales Büro

INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE  
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation 6: <b>C12N 15/11, 5/10, C07K 16/18, 14/47, A61K 48/00, 38/17</b>		A3	(11) Internationale Veröffentlichungsnummer: <b>WO 99/54446</b>  (43) Internationales Veröffentlichungsdatum: <b>28. Oktober 1999 (28.10.99)</b>
(21) Internationales Aktenzeichen: <b>PCT/DE99/01096</b>  (22) Internationales Anmeldedatum: <b>8. April 1999 (08.04.99)</b>		(81) Bestimmungsstaaten: JP, US, europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(30) Prioritätsdaten: <b>198 18 598.7 19. April 1998 (19.04.98) DE</b>		Veröffentlicht <i>Mit internationalem Recherchenbericht.</i>	
(71) Anmelder ( <i>für alle Bestimmungsstaaten ausser US</i> ): <b>METAGEN GESELLSCHAFT FÜR GENOMFORSCHUNG MBH [DE/DE]; Ihnestrasse 63, D-14195 Berlin (DE).</b>		(88) Veröffentlichungsdatum des internationalen Recherchenberichts: <b>2. Juni 2000 (02.06.00)</b>	
(72) Erfinder; und <b>(75) Erfinder/Anmelder (nur für US): SPECHT, Thomas [DE/DE]; Grabenstrasse 14, D-12209 Berlin (DE). HINZMANN, Bernd [DE/DE]; Parkstrasse 19, D-13127 Berlin (DE). SCHMITT, Armin [DE/DE]; Laubacher Strasse 6/II, D-14197 Berlin (DE). PILARSKY, Christian [DE/DE]; Heinrich-Lange-Strasse 13c, D-01474 Schönfeld-Weißenberg (DE). DAHL, Edgar [DE/DE]; Eleonore-Procheska-Strasse 6, D-14480 Potsdam (DE). ROSENTHAL, André [DE/DE]; Koppenplatz 10, D-10115 Berlin (DE).</b>			

**(54) Title:** HUMAN NUCLEIC ACID SEQUENCES WHICH ARE OVEREXPRESSED IN NORMAL PANCREAS TISSUE**(54) Bezeichnung:** MENSCHLICHE NUKLEINSÄURESEQUENZEN, DIE IN PANKREASNORMALGEWEBE ÜBEREXPRIMIERT SIND**(57) Abstract**

The invention relates to human nucleic acid sequences – mRNA, cDNA, genomic sequences – from normal pancreas tissue, coding for the genetic product or parts thereof. The invention also relates to the utilization of said sequences, to the polypeptides obtained through said sequences and their utilization.

**(57) Zusammenfassung**

Es werden menschliche Nukleinsäuresequenzen – mRNA, cDNA, genomische Sequenzen – aus Pankreasnormalgewebe, die für Genprodukte oder Teile davon kodieren, und deren Verwendung beschrieben. Es werden weiterhin die über die Sequenzen erhältlichen Polypeptide und deren Verwendung beschrieben.

### ***LEDIGLICH ZUR INFORMATION***

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidschan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland	ML	Mali	TR	Türkei
BG	Bulgarien	HU	Ungarn	MN	Mongolei	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MR	Mauretanien	UA	Ukraine
BR	Brasilien	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Island	MX	Mexiko	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	NE	Niger	UZ	Usbekistan
CF	Zentralafrikanische Republik	JP	Japan	NL	Niederlande	VN	Vietnam
CG	Kongo	KE	Kenia	NO	Norwegen	YU	Jugoslawien
CH	Schweiz	KG	Kirgisistan	NZ	Neuseeland	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	PL	Polen		
CM	Kamerun	KR	Republik Korea	PT	Portugal		
CN	China	KZ	Kasachstan	RO	Rumänien		
CU	Kuba	LC	St. Lucia	RU	Russische Föderation		
CZ	Tschechische Republik	LI	Liechtenstein	SD	Sudan		
DE	Deutschland	LK	Sri Lanka	SE	Schweden		
DK	Dänemark	LR	Liberia	SG	Singapur		
EE	Estland						

# INTERNATIONAL SEARCH REPORT

Int. Application No  
PCT/DE 99/01096

A. CLASSIFICATION OF SUBJECT MATTER					
IPC 6	C12N15/11	C12N5/10	C07K16/18	C07K14/47	A61K48/00

A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>HOHNE, M.W. ET AL.: "Frequent loss of expression of the potential tumor suppressor gene DCC in ductal pancreatic adenocarcinoma"  <b>CANCER RESEARCH</b>,  vol. 52, no. 9, 1992, pages 2616-2619,  XP002126739  page 2617, first column, lines 15 - 25</p> <p style="text-align: center;">-/-</p>	1

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

### \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*&\* document member of the same patent family

Date of the actual completion of the international search

29 December 1999

Date of mailing of the international search report

12 4 03.00

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.  
Fax: (+31-70) 340-3016

Authorized officer

Alt, G

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/DE 99/01096

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category <sup>2</sup>	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	TAKEDA, J. ET AL.: "A molecular inventory of human pancreatic islets: sequence analysis of 1000 cDNA clones" HUMAN MOLECULAR GENETICS, vol. 2, no. 11, 1993, pages 1793-1798, XP002059393	1
X	the whole document & EMBL Database Entry T10937; Accession No. T10937 12.08.93 Bell, G.I and Takeda, J. Klon hbc153 the whole abstract -----	5,7-10, 34

**INTERNATIONAL SEARCH REPORT**International application No.  
PCT/DE 99 / 01096**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
  
2.  Claims Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3.  Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

See extra sheet

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1 - 38 (partly)

**Remark on Protest**  

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International Application No

PCT/DE 99/01096

1. Claim nos.: 1-38 (in part, insofar as they relate to SEQ ID nos. 14, 68, 69)
2. Claim nos.: 1-38 (in part, insofar as they relate to SEQ ID nos. 24, 39, 40)
3. Claim nos.: 1-38 (in part, insofar as they relate to SEQ ID nos. 25, 41-43)
4. Claim nos.: 1-38 (in part, insofar as they relate to SEQ ID nos. 27 and 44)
5. Claim nos.: 1-38 (in part, insofar as they relate to SEQ ID nos. 28 and 45)
6. Claim nos.: 1-38 (in part, insofar as they relate to SEQ ID nos. 29 and 46)
7. Claim nos.: 1-38 (in part, insofar as they relate to SEQ ID nos. 30 and 47-49)
8. Claim nos.: 1-38 (in part, insofar as they relate to SEQ ID nos. 31 and 50-52)

9. Claim nos.: 1-38 (in part, insofar as they relate to SEQ ID nos. 35 and 56-58)
10. Claim nos.: 1-38 (in part, insofar as they relate to SEQ ID nos. 36 and 59-61)
11. Claim nos.: 1-38 (in part, insofar as they relate to SEQ ID nos. 37, 62, 63)
12. Claim nos.: 1-38 (in part, insofar as they relate to SEQ ID nos. 61, 70 and 71)
13. Claim nos.: 3-22, 28, 29, 33-38 (in part, insofar as they relate to SEQ ID no. 2)
14. Claim nos.: 3-22, 28, 29, 33-38 (in part, insofar as they relate to SEQ ID no. 4)
15. Claim nos.: 3-22, 28, 29, 33-38 (in part, insofar as they relate to SEQ ID no. 6)
16. Claim nos.: 3-22, 28, 29, 33-38 (in part, insofar as they relate to SEQ ID no. 7)
17. Claim nos.: 3-22, 28, 29, 33-38 (in part, insofar as they relate to SEQ ID no. 7)

**INTERNATIONAL SEARCH REPORT**

International Application No

**PCT/DE 99/01096**

18. Claim nos.: 3-22, 28, 29, 33-38 (in part, insofar as they relate to SEQ ID no. 9)
19. Claim nos.: 3-22, 28, 29, 33-38 (in part, insofar as they relate to SEQ ID no. 11)
20. Claim nos.: 3-22, 28, 29, 33-38 (in part, insofar as they relate to SEQ ID no. 12)
21. Claim nos.: 3-22, 28, 29, 33-38 (in part, insofar as they relate to SEQ ID no. 15)
22. Claim nos.: 3-22, 28, 29, 33-38 (in part, insofar as they relate to SEQ ID no. 16)
23. Claim nos.: 3-22, 28, 29, 33-38 (in part, insofar as they relate to SEQ ID no. 17)
24. Claim nos.: 3-22, 28, 29, 33-38 (in part, insofar as they relate to SEQ ID no. 18)
25. Claim nos.: 3-22, 28, 29, 33-38 (in part, insofar as they relate to SEQ ID no. 19)
26. Claim nos.: 3-22, 28, 29, 33-38 (in part, insofar as they relate to SEQ ID no. 21)

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen  
PCT/DE 99/01096

A. KLASSEFIZIERUNG DES ANMELDUNGSGEGENSTANDES  
IPK 6 C12N15/11 C12N5/10 C07K16/18 C07K14/47 A61K48/00  
A61K38/17

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierte Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)  
IPK 6 C12N

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	HOHNE, M.W. ET AL.: "Frequent loss of expression of the potential tumor suppressor gene DCC in ductal pancreatic adenocarcinoma" CANCER RESEARCH, Bd. 52, Nr. 9, 1992, Seiten 2616-2619, XP002126739 Seite 2617, erste Spalte, Zeilen 15 - 25 -/-	1

Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

Siehe Anhang Patentfamilie

- \* Besondere Kategorien von angegebenen Veröffentlichungen :
- \*A\* Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist
- \*E\* Alterses Dokument, das jedoch erst am oder nach dem internationalen Anmeldeatum veröffentlicht worden ist
- \*L\* Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)
- \*O\* Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht
- \*P\* Veröffentlichung, die vor dem internationalen Anmeldeatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

- \*T\* Spätere Veröffentlichung, die nach dem internationalen Anmeldeatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist
- \*X\* Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden
- \*Y\* Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist
- \*&\* Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

Absendedatum des internationalen Recherchenberichts

29. Dezember 1999

12.03.00

Name und Postanschrift der Internationalen Recherchenbehörde  
Europäisches Patentamt, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Alt, G

## INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/DE 99/01096

## C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie <sup>a</sup> :	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	TAKEDA, J. ET AL.: "A molecular inventory of human pancreatic islets: sequence analysis of 1000 cDNA clones" HUMAN MOLECULAR GENETICS, Bd. 2, Nr. 11, 1993, Seiten 1793-1798, XP002059393 das ganze Dokument & EMBL Database Entry T10937; Accession No. T10937 12.08.93 Bell, G.I and Takeda, J. Klon hbc153 das gesamte Abstrakt -----	1
X		5, 7-10, 34

# INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/DE 99/01096

## Feld I Bemerkungen zu den Ansprüchen, die sich als nicht recherchierbar erwiesen haben (Fortsetzung von Punkt 2 auf Blatt 1)

Gemäß Artikel 17(2a) wurde aus folgenden Gründen für bestimmte Ansprüche kein Recherchenbericht erstellt:

1.  Ansprüche Nr. weil sie sich auf Gegenstände beziehen, zu deren Recherche die Behörde nicht verpflichtet ist, nämlich
2.  Ansprüche Nr. weil sie sich auf Teile der internationalen Anmeldung beziehen, die den vorgeschriebenen Anforderungen so wenig entsprechen, daß eine sinnvolle internationale Recherche nicht durchgeführt werden kann, nämlich
3.  Ansprüche Nr. weil es sich dabei um abhängige Ansprüche handelt, die nicht entsprechend Satz 2 und 3 der Regel 6.4 a) abgefaßt sind.

## Feld II Bemerkungen bei mangelnder Einheitlichkeit der Erfindung (Fortsetzung von Punkt 3 auf Blatt 1)

Die internationale Recherchenbehörde hat festgestellt, daß diese internationale Anmeldung mehrere Erfindungen enthält:

Siehe Zusatzblatt

1.  Da der Anmelder alle erforderlichen zusätzlichen Recherchengebühren rechtzeitig entrichtet hat, erstreckt sich dieser internationale Recherchenbericht auf alle recherchierbaren Ansprüche.
2.  Da für alle recherchierbaren Ansprüche die Recherche ohne einen Arbeitsaufwand durchgeführt werden konnte, der eine zusätzliche Recherchengebühr gerechtfertigt hätte, hat die Behörde nicht zur Zahlung einer solchen Gebühr aufgefordert.
3.  Da der Anmelder nur einige der erforderlichen zusätzlichen Recherchengebühren rechtzeitig entrichtet hat, erstreckt sich dieser internationale Recherchenbericht nur auf die Ansprüche, für die Gebühren entrichtet worden sind, nämlich auf die Ansprüche Nr.
4.  Der Anmelder hat die erforderlichen zusätzlichen Recherchengebühren nicht rechtzeitig entrichtet. Der internationale Recherchenbericht beschränkt sich daher auf die in den Ansprüchen zuerst erwähnte Erfindung; diese ist in folgenden Ansprüchen erfaßt:  
1 - 38 (Teilweise)

Bemerkungen hinsichtlich eines Widerspruchs

Die zusätzlichen Gebühren wurden vom Anmelder unter Widerspruch gezahlt.  
 Die Zahlung zusätzlicher Recherchengebühren erfolgte ohne Widerspruch.

WEITERE ANGABEN

PCT/ISA/ 210

1. Ansprüche: 1-38 (teilweise,  
soweit sie sich auf SEQ ID Nos. 14, 68,  
69 beziehen)
  
2. Ansprüche: 1-38 (teilweise,  
soweit sie sich auf SEQ ID Nos. 24, 39,  
40 beziehen)
  
3. Ansprüche: 1-38 (teilweise,  
soweit sie sich auf SEQ ID Nos. 25,  
41-43 beziehen)
  
4. Ansprüche: 1-38 (teilweise,  
soweit sie sich auf SEQ ID Nos. 27 und 44 beziehen  
)
  
5. Ansprüche: 1-38 (teilweise,  
soweit sie sich auf SEQ ID Nos. 28 und 45 beziehen  
)
  
6. Ansprüche: 1-38 (teilweise,  
soweit sie sich auf SEQ ID Nos. 29 und 46 beziehen  
)
  
7. Ansprüche: 1-38 (teilweise,  
soweit sie sich auf SEQ ID Nos. 30 und 47-  
49 beziehen)
  
8. Ansprüche: 1-38 (teilweise,  
soweit sie sich auf SEQ ID Nos. 31 und 50-  
52 beziehen)

WEITERE ANGABEN

PCT/ISA/ 210

9. Ansprüche: 1-38 (teilweise,  
soweit sie sich auf SEQ ID Nos. 35 und 56-  
58 beziehen)

10. Ansprüche: 1-38 (teilweise,  
soweit sie sich auf SEQ ID Nos. 36 und 59-  
61 beziehen)

11. Ansprüche: 1-38 (teilweise,  
soweit sie sich auf SEQ ID Nos. 37, 62,  
63 beziehen)

12. Ansprüche: 1-38 (teilweise,  
soweit sie sich auf SEQ ID Nos. 61, 70,  
71 beziehen)

13. Ansprüche: 3-22, 28, 29, 33-38 (teilweise,  
soweit sie sich auf SEQ ID No. 2 beziehen)

14. Ansprüche: 3-22, 28, 29, 33-38 (teilweise ,  
soweit sie sich auf SEQ ID No. 4 beziehen)

15. Ansprüche: 3-22, 28, 29, 33-38 (teilweise,  
soweit sie sich auf SEQ ID No. 6 beziehen)

16. Ansprüche: 3-22, 28, 29, 33-38 (teilweise,  
soweit sie sich auf SEQ ID No. 7 beziehen)

17. Ansprüche: 3-22, 28, 29, 33-38 (teilweise,  
soweit sie sich auf SEQ ID No. 7 beziehen)

WEITERE ANGABEN	PCT/ISA/ 210
18. Ansprüche: 3-22, 28, 29, 33-38 (teilweise, soweit sie sich auf SEQ ID No. 9 beziehen)	
19. Ansprüche: 3-22, 28, 29, 33-38 (teilweise, soweit sie sich auf SEQ ID No. 11 beziehen)	
20. Ansprüche: 2-33, 28, 29, 33-38 (teilweise, soweit sie sich auf SEQ ID No. 12 beziehen)	
21. Ansprüche: 3-22, 28, 29, 33-38 (teilweise, soweit sie sich auf SEQ ID No. 15 beziehen)	
22. Ansprüche: 3-22, 28, 29, 33-38 (teilweise, soweit sie sich auf SEQ ID No. 16 beziehen)	
23. Ansprüche: 3-22, 28, 29, 33-38 (teilweise, soweit sie sich auf SEQ ID No. 17 beziehen)	
24. Ansprüche: 3-22, 28, 29, 33-38 (teilweise, soweit sie sich auf SEQ ID No. 18 beziehen)	
25. Ansprüche: 3-22, 28, 29, 33-38 (teilweise, soweit sie sich auf SEQ ID No. 19 beziehen)	
26. Ansprüche: 3-22, 28, 29, 33-38 (teilweise, soweit sie sich auf SEQ ID No. 21 beziehen)	